

THE EMERGENCE OF THE
MODERN LANGUAGE SCIENCES
VOLUME 2

THE EMERGENCE OF THE MODERN LANGUAGE SCIENCES

STUDIES ON THE TRANSITION FROM
HISTORICAL-COMPARATIVE TO
STRUCTURAL LINGUISTICS
IN HONOUR OF E. F. K. KOERNER

VOLUME 2: METHODOLOGICAL
PERSPECTIVES AND APPLICATIONS

Edited by

SHEILA EMBLETON

York University

JOHN E. JOSEPH

University of Edinburgh

HANS-JOSEF NIEDEREHE

University of Trier

JOHN BENJAMINS PUBLISHING COMPANY
PHILADELPHIA/AMSTERDAM



The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences — Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

The Editors wish to thank Noam Chomsky for permission to publish the text of his 12 September 1957 letter to Cornelis van Schooneveld in Chapter 18.

Library of Congress Cataloging-in-Publication Data

The emergence of the modern language sciences : studies on the transition from historical-comparative to structural linguistics in honour of E.F.K. Koerner / edited by Sheila Embleton, John E. Joseph, Hans-Josef Niederehe.

v. cm.

Chiefly in English; includes one contribution in French and one in Spanish.

Includes a bibliography of writings by E.F.K. Koerner, 1968-1999

(p.). bibliographical references, and indexes.

Contents: v. 1. Historiographical perspectives -- v. 2. Methodological perspectives and applications.

1. Linguistics--History--19th century. 2. Linguistics--History--20th century. 3. Linguistics--Methodology. 4. Comparative linguistics. 5. Historical linguistics. I. Koerner, E.F.K. II. Embleton, Sheila M. III. Joseph, John Earl. IV. Niederehe, Hans-Josef, 1937-

P73.E47 1999

410'.9'09034--dc21

99-39875

Vol: 2 ISBN 90 272 2188 X (Eur.) / 1 55619 760 8 (US) (Hb., alk. paper)

CIP

© 1999 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. • P.O.Box 75577 • 1070 AN Amsterdam • The Netherlands

John Benjamins North America • P.O.Box 27519 • Philadelphia PA 19118-0519 • USA

Table of Contents

Introduction: Problems of structuralist beginnings (and endings) <i>Sheila Embleton, John E. Joseph and Hans-Josef Niederehe</i>	ix
-------------------------------------------------------------------------------------------------------------------------------------	----

Bibliography of Writings by E. F. K. Koerner, 1968–1999	xxix
---------------------------------------------------------	------

PART IV

Methodological Perspectives

21. Indo-European Methodology, Bloomfield's Central Algonquian, and Sapir's Distant Genetic Relationships <i>Regna Darnell</i>	3
22. The Need for Phonetically Accurate Notation in the General Progress of Linguistics <i>Saul Levin</i>	17
23. The Autonomy of Linguistics: Saussure to Chomsky and beyond <i>Ranko Bugarski</i>	29
24. On the Notion of Zero <i>E. Wyn Roberts</i>	41
25. 'God's Truth' and Structuralism: A new look at an old controversy <i>Gary D. Prideaux</i>	67
26. From Ordered Rules to Ranked Constraints <i>John T. Jensen</i>	81

PART V

Indo-European Linguistics

27. Next of Kin: The search for relatives of Indo-European <i>Allan R. Bomhard</i>	101
---------------------------------------------------------------------------------------	-----

28. Typology and Reconstruction: New trends in comparative historical and diachronic linguistics <i>Thomas V. Gamkrelidze</i>	109
29. Typology and Diachrony of the Middle Voice <i>Helena Kurzová</i>	115
30. Indo-European Numerals Since Szemerényi <i>Carol F. Justus</i>	131
31. Berthold Delbrück and his Contemporaries on 'Tempora' in Sanskrit <i>Vit Bubenik</i>	153

PART VI

Latin and Comparative Romance Linguistics

32. Observations on Two Recently Discovered Latin Inscriptions <i>Philip Baldi</i>	165
33. Comparative, Structural and Sociolinguistic Analyses of the History of the Romance Languages <i>Roger Wright</i>	175
34. Romance Historical Morphology and Empty Affixes <i>Martin Maiden</i>	189
35. Markedness and Morphosyntactic Change Revisited: The case of Romance past participle agreement <i>John Charles Smith</i>	203
36. Romanian and the Balkans: Some comparative perspectives <i>Brian D. Joseph</i>	217

PART VII

Germanic, Caucasian and Asian Linguistics

37. Toward "a Complete Analysis of the Residues": On regular vs. morpholexical approaches to OHG umlaut <i>David J. Holsinger and Joseph C. Salmons</i>	239
38. Narrative Cohesion in the Kensington Runic Text <i>Robert A. Hall, Jr.†</i>	255

TABLE OF CONTENTS

vii

39. The Compound Gerund in Early Modern English <i>Matsuji Tajima</i>	265
40. Gender Affixes in Tsez: Synchrony and diachrony <i>Bernard Comrie</i>	279
41. Once Again on the Reading of the Old Korean 尸 <i>Alexander Vovin</i>	289
42. The Rôle of Historiography in Evaluating the Results of Comparative Linguistic Work: A case study <i>Paul Sidwell</i>	301
Koerner Tabula Gratulatoria	313
Name Index	315
Subject Index	321
Contents of Volume One	333



E. F. K. Koerner

Ottawa, 20 November 1998 — Gustave Pellerin, Photographer

Introduction

Problems of Structuralist Beginnings (and Endings)

Sheila Embleton, John E. Joseph & Hans-Josef Niederehe

1. The crucial question of Saussure's continuity with linguistic tradition

If some future age reduces the linguistics of the 20th century down to a single word, *structuralism* will surely be it, however *passé* the term may seem now as the century draws to a close. All but the youngest of those who called themselves structuralists during that movement's heyday — roughly, the early 1930s through the early 1970s — are approaching retirement age. Yet self-proclaimed 'New Structuralists' are afoot in the land (see Lieb 1992), and as our historical perspective lengthens, it looks increasingly as though the proclaimed death of structuralism in linguistics was somewhat exaggerated. The 'Chomskyan revolution' involved a clear break at the rhetorical and sociological levels, but a rather murkier one when it comes to the conceptual and the methodological. It is already possible to imagine a not-too-distant generation reckoning that structuralism, in various guises, outlasted the 20th century.

Our understanding of the movement's beginnings has similarly expanded backward beyond the 20th century. When structuralism became a general intellectual movement in Paris in the 1950s, a received view developed that it had originated in the *Cours de linguistique générale* of Ferdinand de Saussure (1857–1913), assembled and published posthumously in 1916. The lectures on which the book were based were given in the years just before the First World War, the same years in which modernism was transforming the arts, and Max Planck (1858–1947) and Albert Einstein (1879–1955) were rewriting the laws of

physics. The War itself brought the old order in Europe to an end. It was easy enough to imagine Saussure's lectures as being of a piece with these other developments, sweeping away whole traditions in the study of language and establishing utterly new and modern modes of linguistic thought. The Saussurean notion that the linguistic sign is constituted by pure difference seemed to have affinities with serialism in music, cubism in art, and relativity in science, and to mark an absolute break with the classical modes that had held sway since the Renaissance. What is more, Saussure's distinction between synchronic and diachronic linguistics was read by later structuralists as a rejection of the historical study of language which, from its nerve centers in Leipzig and Berlin, had so thoroughly dominated linguistics up to the Great War and even after. In undoing the belief that the reconstruction of lost ancestral tongues and historically-based typology were the only scientifically valid ends of linguistics, Saussure released linguists from their bondage to the past. This appeared to align with yet another essential modernist impulse.

This image of Saussure as a revolutionary iconoclast, and of structuralism as bringing the great tradition of linguistics to its knees, is also the one found in Noam Chomsky's (b.1928) *Cartesian Linguistics* of 1966. Chomsky depicts himself as restoring a Renaissance tradition of the analysis of language and thought which the structuralists had betrayed. The dilettantish *Cartesian Linguistics* proved one of the great debacles of Chomsky's professional career, as historians of ideas and historically informed linguists joined forces to make clear to one and all that while facts may not count for much in certain modes of syntactic analysis, historians are a bit less easily persuaded to ignore them. For Chomsky, this simply means that history isn't sufficiently theoretical to be taken seriously, but for many intelligent people who read *Cartesian Linguistics* as the first non-technical book on language and thought by the man being touted as the new guru on those subjects, these reviews made all too clear the limits to his analytic capacities and his own inability to recognize them. In the process, the history of linguistics garnered wider attention among linguists than it had enjoyed for decades.

The ensuing period coincided with, in France, the final apogee of the structuralism that had been extrapolated out of linguistics by Claude Lévi-Strauss (b.1907) and others, and the birth of post-structuralism with the early works of Jacques Derrida (b.1930); and in North America and Britain, the first wave of imported French structuralism. So it is not surprising that Saussure and the origins of structuralism were popular topics for historical inquiry and speculation. Among at least some of the French structuralists there was particular interest in linking Saussure with Karl Marx (1818–1883). The links were lent apparent

weight by Saussure's central concept of 'value' and the fact that Marx's impersonal, unconscious economic forces, joint issue of Adam Smith's (1723–1790) 'invisible hand' and G. W. F. Hegel's (1770–1831) Romantic constructs of history and human action, suggested a rough equation with Saussure's unconscious, socially shared *langue*.

Jonathan Culler's *Ferdinand de Saussure* (1975) focussed instead on Saussure's links to his close contemporaries Sigmund Freud (1856–1939) and Émile Durkheim (1858–1917). Although Saussure's son Raymond would from the 1920s become one of France's leading Freudian psychologists, the possibility is remote that Saussure might have known Freud's early writings well enough to have been in any way engaged with them, let alone influenced by them. However, Durkheim is another, more complicated matter. It was well known that Saussure's student and friend Antoine Meillet (1866–1936) was the principal linguistic contributor to Durkheim's journal *L'Année sociologique*. Moreover, in two well-known articles of 1933, Witold Doroszewski (1905–1976) had cited second-hand information confirming that Saussure had taken from Durkheim his notion of *langue* as a 'social fact'. Hans Aarsleff, who had been one of the most scathing and effective critics of Chomsky's *Cartesian Linguistics* (Aarsleff 1970), would take the Saussure-Durkheim link a couple of steps further, first by asserting that the debate in the press between Durkheim and Gabriel de Tarde (1843–1904) could not possibly have escaped Saussure's attention, and moreover that other key Saussurean concepts could be traced to the work of the widely read writer Hippolyte Taine (1828–1893) (see the collection of papers published as Aarsleff 1982).

While literary theorists like Culler and historians of ideas like Aarsleff were looking for the sources of Saussure's vision of language among some of the most prominent non-linguistic thinkers of his time, some linguists were discovering influences from within the field. Eugenio Coseriu (b.1921) made the case that many of the concepts for which Saussure received credit had in fact originated in the work of Georg von der Gabelentz (1840–1893) (see Coseriu 1967; Gabelentz 1901). Tullio De Mauro, while questioning the link to Gabelentz, raised a number of additional possible sources. Taking a cue from Collinder (1962), he gave particular attention to links between the thought of Saussure and that of the Swedish linguist Adolf Noreen (1854–1925) (see Saussure 1973 [1916]: 390–394, 396), and he equates one passage from an unpublished manuscript of Saussure's with a passage from Hegel's *Encyclopedia* (ibid. 362). Roman Jakobson (1896–1982) pointed to a number of Saussurean precursors including fellow Russians like Filipp Fedorovich Fortunatov (1848–1914), as well as the 'Hegelian' (in Jakobson's view) Victor Henry (1850–1907), who had

in fact been part of Saussure's circle during the latter's Parisian years, and whose book *Antinomies linguistiques* (1896) foreshadows certain philosophical themes that re-emerge in Saussure's *Cours* (see further Joseph 1996a). Again, links to Hegel would be interpreted by many as indirect links to Marx, given the extent to which Marx's conception of history is fundamentally Hegelian.

The 'influences' mentioned are actually just a few of the many proposed in the late 1960s and early 70s, so many that it no longer seemed clear whether Saussure had actually come up with *any* ideas himself, or was simply an intellectual sponge soaking up the ideas around him and letting them drip out again without acknowledging their sources. It is worthy of note that those contributing to this state of affairs included some extremely prominent figures from three generations of linguistic-cum-literary studies, Jakobson, Coseriu, and Culler; De Mauro, the scholar entrusted by Saussure's publisher with the task of editing the first critical edition of the *Cours de linguistique générale*; and Aarsleff, who had gained a substantial reputation on the basis of his demolition of Chomsky's *Cartesian Linguistics* in the pages of *Language*.

When articles, then a book, challenging some of these ideas began appearing from the pen of an ambitious Ph.D. student at Simon Fraser University, they were received in the typical fashion by the established scholars — they tried to crush them under their heel. Reading the reactions to the young Koerner's criticisms of them in the notes to their works, it seems that neither Coseriu nor Aarsleff had any inkling of the sort of character they were dealing with. Most young scholars would never recover from such blistering, *ad hominem* assaults by established masters. Koerner, however, stood his ground. Against the massive force of all the views concerning Saussure and the beginnings and later history of structuralism sketched in the preceding paragraphs, Koerner established a crucial point which would prove to have far-ranging consequences. It was that, far from being an iconoclastic revolutionary, *Saussure had remained essentially true to his historically-based Neogrammarian training.*

2. Koernerian themes and variations

On the surface, this is quite a conservative interpretation of Saussure, certainly far less exciting than the various attempts to link him to Marx and Freud. In the heat of the worldwide structuralist movement, the desire to elevate Saussure to a status comparable to that of these great thinkers was such that normal rules of evidence were ecstatically suspended. And those who were making their own reputations with speculative interpretations of Saussure were bound to look

askance upon a young upstart spoiling the party by questioning their validity. The fact that he was doing this in such a straightforward, commonsense way, as against the sometimes lofty flights of rhetoric they were used to deploying, led some of them into the deadliest trap of all: underestimating their opponent. To this day, Koerner's achievement as well as his individual pieces of scholarship are underrated by those of his contemporaries who persist in the opinion that, in modern scholarship, excitement is preferable to solidity, and obscurantism to clarity. Happily, this view is now rarer among younger scholars than among their teachers; fans of Koerner's work are likelier to have been born after the publication of *Cartesian Linguistics* than before it.

Among the implications that fall out from Koerner's revisionist view of Saussure can be traced many of the themes that have dominated his work since the early 70s:

- The origins of structuralism can be found mainly in developments *within* the linguistics of the second half (and particularly the final third) of the 19th century, rather than outside it.
- In spite of the compelling affinities, there is no proof of Saussure having been influenced by Durkheim, and even Doroszewski's testimony needed to be called into question.
- In no way did Saussure see his advocacy of synchronic linguistics as any kind of rejection of diachronic linguistics; rather synchrony and diachrony together were integral to the methodological improvements he wished to bring to traditional historical inquiry.
- Gabelentz, an outsider to the Neogrammarian mainstream, is by all evidence a less convincing source of Saussurean ideas than Hermann Paul (1846–1921), whose 1880 *Principien der Sprachgeschichte* (later editions spelled '*Prinzipien*') was regarded as a 'bible' by the Neogrammarians who trained Saussure, and William Dwight Whitney (1827–1894), the American linguist whom they likewise revered (and the question of whose influence on Saussure is discussed at length by De Mauro in Saussure 1973 [1916]: 332–334, 360–364, 387–388; cf. Joseph 1988).
- Insufficient attention has been given to a number of figures in 19th century German linguistics who prefigure structuralism, more than a few of whom can be characterized as having been influenced by the natural sciences, though not always in the way such influence is often depicted, a good example being the evolutionary interests of August Schleicher (1821–1868).
- Significant parallels exist among various 'revolutions' proclaimed in modern linguistics, from Sir William Jones's (1746–1794) 'discovery' of Indo-European linguistic affinities in 1786, to Franz Bopp's (1791–1867)

methodological refinements of 1816, to the Neogrammarian 'revolution' of 1876, and onward to Saussure in 1916, Jakobson in 1929, Chomsky in 1957.

- The version of the history of structuralism promulgated by Jakobson, although he himself is key to that history, is on certain points self-serving and dubious.
- The version of the end of structuralism promulgated by Chomsky is likewise self-serving and dubious, particularly in view of the direct continuity between Chomsky's early methodology and that of structuralists like Jakobson, Zellig Harris (1909–1992), and Charles Hockett (b.1916).
- Significant 19th-to-20th century continuity also existed in other branches of modern linguistics which assign themselves purely contemporary origins, for example the line of anthropological linguistics which extends from Franz Boas (1859–1942) through Edward Sapir (1884–1939) to Dell Hymes (b.1927), and which in some respects continues a 'Humboldtian trend'.
- The entire question of 'influence' in the history of linguistics is problematical, along with much else in the methodology and metalanguage of the undertaking.

At least two things should be obvious from this list. First, these challenges to established views were bold to the point of audacity. That the man who made them should have accrued many enemies is not surprising; and if it is true that one's impact can be judged by the importance of one's enemies, few living scholars could compete. Second, although it is far from the case that all the points listed are now universally accepted without contention, enough of them have come to be accepted by enough people that it can be stated without fear of exaggeration that they have amounted to a rewriting of the history of modern linguistics. Perhaps less obviously, but no less importantly, they have helped to undo the rhetorical strategies by which a narrow range of approaches within contemporary linguistics became and remained so dominant for several decades that alternative and potentially more useful approaches were stifled.

Our admiration for Koerner's work is not uncritical, even if we might be forgiven for pretending it were so in the present context. In fact, one of the editors of this volume has taken issue with Koerner on a number of points, some of them going straight to the heart of the program outlined above: see Joseph (1989) on the ongoing need for a definitive statement of the case for Paul's influence on Saussure; Joseph (1991) and (1995) on the problems with his implicit criteria for 'revolutions'; Joseph (1996) for some caution on the 'Humboldtian trend'. But it is worthy of note that the two most recent of these have appeared, undiluted, in Koerner's own journal, and that over the years his work has steadily taken account of criticisms made of it, and gained accordingly in argumentative force.

It would be a complex task indeed to place Koerner in the kind of intellectual lineage he is fond of tracing for others. The lineage would have to be triangulated across the three countries in which we three editors reside, namely Germany, Britain, and Canada, none of which however is his native land. Ernst Frideryk Konrad Koerner was born on 5 February 1939 in Mlewiec near Torun (Thorn), Poland, and his childhood encompassed the brutal aftermath of the Second World War. From 1960 to 1962 he served in the German air force, staying in the reserves until 1968, a year after his promotion to the rank of Captain. From 1962 to 1965 he studied English and German philology, the history of art, pedagogy, philosophy, and applied linguistics at the Universities of Göttingen, Berlin, and Edinburgh, and completed a B.Phil. at the Freie Universität Berlin in 1965. In 1965–66 he taught German and English at the Collège Notre Dame in Valenciennes, France, while also doing some German teaching for the Goethe Institutes of Lille and Brussels. From 1966 to 1968 he studied German and English philology at the Justus Liebig University Giessen, obtaining the state diploma for high-school teaching and M.A. in 1968.

From there he went to Simon Fraser University in Burnaby/Vancouver, British Columbia, where he completed his Ph.D. in 1971. He spent the next few years occupying posts as research associate, at the University of Texas at Austin (1972), the University of Indiana (1972–73), and the German Research Foundation attached to the University of Regensburg, Germany (1973–75), after which he was appointed as Lecturer in the Institute of General and Indo-European Linguistics at the University of Regensburg and, concurrently, Research Fellow in the Fritz Thyssen Foundation. Finally, in 1976 he took up the position as Associate Professor, then Professor of Linguistics, and (until 1979) Director of the Linguistics Documentation Centre of the University of Ottawa, from which he is retiring in 1999. During the Ottawa years he has also held Visiting Professorships at several leading universities in the U.S., Europe, Asia, and South America. In 1994 he was awarded an honorary doctorate by the University of Sofia, and in 1995 he received a diploma and a medal of merit from the Nicholas Copernicus University of Torun, the town of his birth. In 1997 he was elected as a Fellow of the Royal Society of Canada, and in 1998 as a Fellow of the Royal Society for the Arts (London). He has been the recipient of two previous volumes in his honor, the first a personal bibliography with accompanying tributes from colleagues on the occasion of his fiftieth birthday (Cowan & Foster eds. 1989), the second a recent volume entitled *Professing Koernerian Linguistics*, published by the Association for the History of Language centered at the University of Melbourne (Kirk & Sidwell eds. 1998).

Although it was in Canada that he completed his PhD and would spend

most of his academic life, the fact that his thesis director at Simon Fraser University was Geoffrey L. Bursill-Hall (1920–1998) links him directly with the London School. Bursill-Hall had done his own doctoral work at the University of London under Robert Henry Robins (b.1921), who himself, with the support of his mentor John Rupert Firth (1890–1960), had played an important role in reviving interest in the history of linguistics from the 1950s onward. As a specialist in medieval speculative grammar, Bursill-Hall might not have been the first person one would have predicted to supervise what turned out to be an important thesis on the sources of Saussure's linguistic ideas. But of course what he gave his student was first and foremost the sound methodological grounding which has sustained all of Koerner's subsequent endeavors, including both his own research and his creation and maintenance of avenues for the research of others.

The list of his publications that follows this Introduction will show the extent of his research productivity since 1968 — it includes ten books authored, 32 books edited or co-edited, 12 chapters in books or collective volumes, 55 articles in refereed journals, and 46 further papers in publications of various sorts — but to this must be added the enormous proportion of his research energies that have gone into helping shape the publications of other scholars contributing to the journals and book series he has founded. These will be discussed in the following section.

3. Koerner's impact as organizer and editor

From the beginning of his time as a Fellow of the German Research Foundation in Regensburg, Koerner started organizing various publication projects centered on the history of linguistics. He planned things on a large scale and, at least for the German university system of that time, brought them out at an incredibly fast pace. When he published his dissertation on Ferdinand de Saussure in 1973, he announced at the same time, on the cover page of the dissertation, the creation of the 'sole journal for the history of linguistics', *Historiographia Linguistica*, now subtitled *International Journal for the History of the Language Sciences*. It was singlehandedly planned and prepared by him, with the technical help of John Benjamins, then a little-known antiquarian bookseller in Amsterdam whom Koerner persuaded to enter the publishing business by reprinting some books on language which he had in stock and which Koerner had had difficulty locating in various academic libraries around the world. When launching *Historiographia Linguistica* Koerner also announced the accompanying series *Amsterdam Studies in the Theory of Linguistic Sciences*, whose first volume was written by Koerner

himself, on *The Importance of F. Techmer's Internationale Zeitschrift für Allgemeine Sprachwissenschaft*, which in retrospect might be seen as a sort of *apologia* for the enterprise Koerner himself was launching.

Apologies, or at least justifications, were in order: for these activities did not meet with the general approval of the 'full professors' of Regensburg University – hard as it may be to understand today, unless one were to attribute to those distinguished scholars motives unworthy of their rank. In any case, their reaction formed the background to Koerner's move to North America after what was initially a highly promising start in the German university system, taking his ambitious projects and plans with him.

Meanwhile, things had started moving in the field of the history of the language sciences. Koerner's account of the new activities reads as follows: "In Spring 1975, when a 'Kolloquium zur Wissenschaftsgeschichte der Romanistik' was announced for October 1975 at the University of Trier [organized by one of the editors of this volume (Niederehe), who in 1976 was invited by Koerner to become associate editor of *Historiographia Linguistica*], and an 'International Conference on Medieval Grammar' was scheduled to take place in February 1976 at the University of California at Davis, I boldly wrote that an 'International Symposium on the History of Linguistics' was being considered for the Summer of 1976, adding that this was 'the logical outgrowth of the increasing interest in the subject [which had] become more and more evident in recent years'" (Koerner 1980: xi). But these optimistic plans of his were realized only in 1978, when he organized an international conference in 1978 at the University of Ottawa, on the History of the Language Sciences. This was a huge success, and led to the establishment of a regular series of these triennial International Conferences on the History of the Language Sciences (ICHoLS), of which the founding one in Ottawa retroactively became the first.

The Ottawa ICHoLS marked the start of the 'globalization' of the emerging discipline. That same year, the Société d'Histoire et d'Épistémologie des Sciences du Langage (SHESL) was founded in France, and this was followed by the foundation of a number of national historiographic societies whose history would deserve a sketch of its own. In most of the cases, Koerner was behind the scenes planting ideas, offering advice, and perhaps most importantly, keeping the momentum going whenever it threatened to flag. In one case only did he take a direct role in the foundation of one of these societies, the North American Association for the History of the Language Sciences (NAAHoLS, which meets annually with the Linguistic Society of America). His indirect yet important role in helping the Henry Sweet Society for the Study of Linguistic Ideas in the U.K. has recently been documented (Koerner 1998h).

Koerner's talent for organizing and his capacity for sheer hard work revolutionized the field by providing an international network for scholars from all over the world to communicate their work to one another. The effects of this became all the greater as the establishment of the international organization gave an impetus for national organizations to be established. Koerner has also lent his time and organizing talents unstintingly in other ways, for example furthering the cause of the history of linguistics in the last few years by making several trips to Brazil to help scholars in that country set up programs and organizations in the history of linguistics — and not just to escape the rigors of the Ottawa winter!

Koerner has also been a major contributor of organizational and editorial talent to historical linguistics. We would like to single out here his establishment in 1984 of *Diachronica: International Journal for Historical Linguistics*, first published by Georg Olms and now by John Benjamins. *Diachronica* provides a forum for the presentation and discussion of all aspects of language change, in any and all languages of the globe, seeking particularly contributions which combine theoretical interest and philological data-oriented acumen. Until December 1998, Koerner was General Editor, handling virtually all editorial correspondence with authors and referees, and preparing all camera-ready copy for the two issues of 150–200 pages each per year. The latter is no mean feat, given the profusion of word-processing programs (both Mac and IBM) and exotic fonts that could be found in the incoming manuscripts.

Diachronica has a special importance in another way too. It was established at a time when historical linguistics was out of fashion in North America, and it was therefore very difficult even for very good work in the field to get published (Europeans had a tendency to publish only their own work — or at least it seemed that way to many North Americans — and North American journals tended not to publish historical linguistics). Many linguists likely owe their careers to the existence of this journal. *Diachronica* was also instrumental in breaking historical linguistics out of the mold of being seen as virtually coterminous with Indo-European linguistics, and in bringing historical research in non-Indo-European languages (formerly largely published in specialist area journals) to the attention of Indo-European and general historical linguists. The importance of this widening out of the field cannot be overestimated. Another valuable contribution to historical linguistics, which deserves to be singled out from among the others, has been making his book series *Current Issues in Linguistic Theory* (CILT, on which see below) into the 'custodian' of our collective memories of the regularly-held International Conferences on Historical Linguistics (ICHL) and International Conferences on English Historical Linguistics (ICEHL). Not only has CILT published the proceedings, but Koerner himself has

more than once stepped in as the editor (sometimes with his own byline, sometimes purely as a ghost-editor, with no official recognition at all), when the official editor has either abandoned the project or proven too tardy or otherwise incompetent. This has ensured the maintenance of continuity, timeliness, and standards.

Koerner's editorial and organizational skills and dynamism have affected the entire field of linguistics as well in yet further ways. In 1975, he established a book series, again published by John Benjamins, called *Current Issues in Linguistic Theory* (CILT), whose name is self-explanatory. Nearly 200 volumes in this series have now appeared — a remarkable expenditure of time and energy in reading, commenting, editing, overseeing each author's production of camera-ready copy, all with no staff assistance. This series was established at a time when it was very difficult for linguistic research falling outside of (some variant of) the dominant Chomskyan paradigm to be published, especially in North America. Not only do many linguists owe their careers (as above) to the existence of this series, but linguistics owes at least some of its current richness to the fact that somebody took the initiative to preserve the pluralism of the field's voices, and to prevent the field from degenerating into a single party-line orthodoxy. Koerner's other book series, some of which were already mentioned above, have also had major and lasting impacts on their subfields: *Studies in the History of the Language Sciences*, established in 1973, is nearing its 100th volume; *Amsterdam Classics in Linguistics, 1800–1925*, established in 1974, is nearing 20 volumes; *Classics in Psycholinguistics*, established in 1978, has now published 5 volumes; the extremely useful *Library and Information Sources in Linguistics*, established in 1977, has now passed the 25-volume mark.

Koerner has set new standards of scholarship in various fields of linguistics, beginning with his own work in the historiography of linguistics, in terms of depth and breadth of research, attention to detail, accuracy, and impartiality. This is partly because his own standards have been emulated by others, but also because he has been the founder and moving force behind many highly respected book series in this particular field, its most widely read journal, and more recently *The Concise History of the Language Sciences: From the Sumerians to the Cognitivists* (with R. E. Asher, published by Pergamon Press in 1995). Most recently, he has been working (with Sylvain Auroux, Hans-Josef Niederehe, and Kees Versteegh) on the massive volume *Geschichte der Sprachwissenschaften/History of the Language Sciences/Histoire des sciences du langage* for Walter de Gruyter's HSK [Handbücher zur Sprach- und Kommunikationswissenschaft] series — another enormous editorial undertaking. All of this activity has vastly increased the visibility of this particular field, and forced all linguists, whatever

their area, to be at least a bit more attentive to the history of their own discipline. But it has also enabled Koerner, through his extremely active editorial roles, to maintain and enhance rigorous standards for others to abide by, thus virtually single-handedly improving the quality of scholarship in the entire field.

4. Introduction to the papers in the present volume

In planning these volumes the editors faced a dilemma: Koerner has had an impact upon so many areas of the study of language that a *Festschrift* including all these areas would be extremely diffuse, in a time when it is well known that university librarians throughout the world have dropped such volumes to the bottom of their list of purchasing priorities. We decided that Koerner would be better honored with a relatively coherent collection organized around a central theme having to do with the major thrust of his own research and editorial work in the history of the language sciences and historical linguistics. Such a collection would make a greater intellectual contribution to the field and (as the publisher gently reminded us) would find a wider distribution and availability. But how hard this decision was to make will be evident from the absence of certain specialists in the history of pre-19th century linguistics who have had long and fruitful associations with the honoree. Some of these specialists were able to take up our invitation to extend their historical perspective into the prescribed time frame – with some quite interesting results, such as Joseph L. Subbiondo's look at how 17th-century British linguistics, his own research area, is treated in the work of the great 20th-century British linguist J. R. Firth – others understandably felt that such a leap forward from ancient Rome or the early Renaissance was beyond the responsible limits of their particular research programs. We repeat our apologies to them, and express our regret to the honoree that not everyone could be accommodated between these covers who ought, in an ideal world, to have been represented in a *Festschrift* for him.

We should also give special mention to another scholar and friend absent from these pages, Paul B. Salmon (1921–1997), who agreed to contribute but died before undertaking the writing of his paper. Another great scholar, Robert A. Hall, Jr. (1911–1997), is likewise no longer with us, but fortunately was able to complete his contribution before his death.

The 42 papers here assembled are grouped into seven sections, arranged symmetrically with three historiographical parts and three historical-linguistic parts surrounding a single set of methodologically-oriented papers. The papers in the historiographical sections are focused on the work of particular linguists of

the 19th and 20th centuries. Those in the methodological section look at particular concepts and procedures across the work of several individuals. Finally, the sections on historical linguistics concern the analysis of particular phenomena in particular languages or language families, in many cases comparing current views with those of the late 19th century or examining how the former developed out of the latter.

Volume one: Historiographical Perspectives

Part One, "Before Saussure"

Begins with **Lia Formigari** tracing the philosophical heritage of 19th-century linguistics from the two centuries preceding it. This is a topic that has figured prominently in Koerner's many articles on Schlegel and other key figures at the start of modern linguistics. Following Formigari's erudite overview is a set of papers discussing four relatively little-known linguists of the mid-19th century, two of whom made enduring contributions to syntax and phonology. The Dutch linguist Matthias de Vries is the subject of **Jan Noordegraaf's** paper, which looks at the 17th-century sources behind his work as well as at how he was influenced by another of Koerner's favorite subjects, August Schleicher. **Gerda Hassler** examines Pierre Benjamin Lafaye's 1841 *Dictionnaire des synonymes* and the influence it may have had on the semantic theories formulated later in the century by Carl Abel. **D. Gary Miller** considers the treatment of syntax in Samuel Kleinschmidt's classic 1851 grammar of Greenlandic in the light of current theory, and **John E. Joseph** furnishes new information concerning A. Dufriche-Desgenettes, inventor of the term *phonème* and the subject of a 1976 paper by Koerner. The last two papers of the opening section examine the work of three rough contemporaries of Saussure, with **Kurt R. Jankowsky** taking up the founding roles played by Henry Sweet and Eduard Sievers in establishing modern phonetics, and **Maxim I. Stamenov** surveying the work of Ivan Georgov in establishing first-language acquisition as a field of inquiry in the first decade of the 20th century.

Part Two, "Saussure"

Focuses on the figure who stands at both the head and the heart of Koerner's personal research program. The first two papers of this section attempt to place Saussure's ideas in the broader linguistic context of the last quarter of the 19th and first decades of the 20th centuries. **Douglas A. Kibbee** considers the concept of the *masse parlante* in Saussure's *Cours*, and how it relates to the concept of 'the people' in the work of his predecessors and contemporaries. **George Wolf**

analyzes Saussure's own views of his 19th-century forerunners Bopp, Schlegel, and the Neogrammarians, noting how the version popularized in the *Cours* misrepresents what is in the source materials. The next two papers take up some of the core concerns of the *Cours* and question how they may or may not contrast with other approaches, in one case that taken by Saussure himself in his notorious analysis of anagrams in poetry, and the other that taken by one of the greatest linguistic philosophers of the 20th century, Ludwig Wittgenstein. In their joint paper, **W. Terrence Gordon and Henry G. Schogt** challenge the widely-held notion that Saussure's obsessive search for scrambled messages hidden in ancient and modern poems imply a deep contradiction with the linguistic principles espoused in the *Cours*. **Brigitte Nerlich** tries to determine what Saussure might have meant by 'identity', a concept which stands at the very foundation of *langue*, and draws enlightening parallels with Wittgenstein's direct reflections on what makes words or other linguistic units 'the same'. The last two papers of this section contribute considerable new material to Saussurean studies, one by presenting unpublished manuscript material by Saussure and the other by tracing his influence in a country where the Saussurean heritage has heretofore gone unexamined. **Anders Ahlqvist** introduces Saussure's Old Irish copybook and elucidates its often elliptical entries, while also providing background information on both earlier and later developments in the study of Old Irish. **Jivco Boyadjiev's** survey of Saussure's reception in Bulgaria is simultaneously a capsule history of the development of modern linguistics in that country (which, by the way, was the first to award Koerner with an honorary doctorate).

Part Three, "Historiographical Perspectives III: After Saussure", is actually continuous with Part Four, "Methodological Perspectives", since all of the latter are directly concerned with post-Saussurean developments. Part Three begins with three papers on linguists in Spain, Germany, and Britain in the first half of the 20th century who did not necessarily see themselves as Saussurean but whose work has important parallels with early structuralism. **Emilio Ridruejo** focuses on how the conception of phonetic change in the work of the great Hispanist and linguist Ramón Menéndez Pidal developed from 1926 onwards in a post-Neogrammarian vein with clear links to structuralism. He examines the question of Menéndez Pidal's relationship to idealism, and draws interesting links to Koerner's research on the history of sociolinguistics which has shown its deep indebtedness to a number of Romance linguists trained or influenced by Menéndez Pidal. **Werner Hüllen** provides an overview of the nationalistic themes that emerged in German linguistics in the period 1914–1945, a topic which is nowadays coming, somewhat belatedly, to the forefront of the

research agenda, including in Koerner's most recent work on ideological trends in German historical linguistics of the period. **Joseph L. Subbiondo** analyzes the treatment of linguistic predecessors in the 1937 book *The Tongues of Men* by J. R. Firth, whose complex relationship to Saussure has also been the subject of recent work by Koerner.

Shifting gears and continents, the next paper, by **Cristina Altman**, is a history of the reception of structuralist linguistics in Brazil, a country which in fact played a cardinal role in the development of structuralism as the setting for Claude Lévi-Strauss's 1955 book *Tristes tropiques*, which thrust structuralist ideas forward from linguistics onto the general intellectual scene. The following two papers concern (and the third indirectly concerns) the figure who so radically transformed structuralism, Chomsky, the subject of a series of papers by Koerner examining how 'revolutionary' his linguistics has been. With characteristic irreverence, **Stephen O. Murray** punctures yet another Chomskyan myth, providing new documentary evidence contradicting Chomsky's oft-repeated story of how no one would publish his doctoral thesis in the 1950s or 60s. **Danny D. Steinberg** goes back to Chomsky's early writings to show that prior to 1959 he was an anti-mentalistic formalist, and didn't fully embrace mentalism until 1965, when he did so in a way that, in Steinberg's view, rendered his theory of grammar useless for psycholinguistic inquiry. Since Steinberg has for many years now been a leading figure in psycholinguistics in Japan, his paper links interestingly with the one that closes this section, in which **Joseph F. Kess and Tadao Miyamoto** trace the origins of Japanese psycholinguistics to, not linguistics, but psychology, going back to the early decades of the 20th century.

Volume two: Methodological Perspectives and Applications

Part Four, "Methodological Perspectives", opens with a paper by **Regna Darnell** that serves as a bridge between the historiographical sections and what will follow. After exposing the essential tension between the linguistic and anthropological approaches in contemporary Americanist linguistics (the history of which has been one of Koerner's enduring interests), Darnell conducts a historical search for its origins, and ends up focusing on a set of contrasts in the work of Bloomfield and Sapir (again both regular Koernerian subjects). In so doing she demonstrates the power of historiographical inquiry in providing solutions to present-day problems on the grand scale. **Saul Levin's** paper likewise raises a persistent issue, this time on the micro rather than the macro

level: the need for phonetically accurate notation. Levin points out how much of the progress of Western linguistics over the last 2500 years has hinged precisely on the zeal for orthography in the literal sense. **Ranko Bugarski** enquires into the development of 'autonomy' as a primary goal of linguistics, arguing that even if it has been a necessary goal, it need not and should not mean the cutting off of linguistics from adjacent disciplines. The next two papers take up leading ideas of the structuralist period and bring them forward into today's context. **E. Wyn Roberts** offers an encyclopedic history of the concept of 'zero' in linguistics from Saussure to the Bloomfieldians, the Prague School, generative phonology, and many others, including key figures in British phonology. **Gary D. Prideaux** examines the structuralist-derived conceptions of constituent structure and markedness in the light of the 'God's truth' vs. 'hocus pocus' controversy of the Bloomfieldians. He argues on the basis of original empirical research presented here that both constituent structure and markedness must be God's truths and not analytical fictions, and in the process he provides some useful perspectives about a controversy which, contrary to popular belief, is far from dead. The section closes with **John T. Jensen's** account of the move from ordered rules to ranked constraints in work since the 1960s, culminating with an outline of optimality theory including one of its most recent developments, sympathy theory.

Part Five, "Indo-European Linguistics", is the first of the three sections devoted to historical linguistic inquiry. The connection of this section to Koerner's own work is of course that nearly all the major 19th-century linguists whom he has studied, including Saussure, were Indo-Europeanists. The first two papers are about the reconstruction of 'Nostratic' and Indo-European and have been contributed by leading figures in those fields. **Allan R. Bomhard** recounts the history of the search for the macro-family called Nostratic, from the mid-19th century onward to its resuscitation from the mid-1960s to the present. **Thomas V. Gamkrelidze** provides a concise and accessible introduction to the Glottalic Theory which he formulated in work with V. V. Ivanov, placing it usefully into historical perspective. **Helena Kurzová** considers the treatment of middles in Indo-European comparative linguistics from the late 19th century to the present, and **Carol F. Justus** argues for a recomplexification of the Proto-Indo-European numeral system in the light of theoretical and empirical advances since Szemerényi's influential 1960 claim that it must have been a decimal system, in connection with his etymological analysis of the IE word for ten, **dék'mt*, as deriving from **de-k'ont* 'two hands'. Completing this section is **Vit Bubenik's** look at how the Sanskrit tense and aspect system was analyzed by Delbrück, Whitney and the less well-known J. S. Speijer, who published the first complete treatment of Classical Sanskrit syntax in 1886.

Part Six, "Latin and Comparative Romance Linguistics", covers an area which, like the preceding, has benefitted greatly in recent decades from the new research outlets to which Koerner's efforts have given rise – his Current Issues in Linguistic Theory series, for instance, has published most of the recent proceedings volumes of the Linguistic Symposium on Romance Languages, as well as many other important collections and monographs on Romance linguistics. The field has also profited from the improved understanding of its 19th-century roots made possible by Koerner's own research. The section starts off with **Philip Baldi's** observations on the significance of two recently discovered Latin inscriptions for the reconstruction of Proto-Italic. **Roger Wright** considers how the structuralist notion of the language 'system' distorted work in Romance historical linguistics through much of the 20th century, and how a historical sociolinguistic perspective, taking account of the nature of bilingualism and variation in existing language communities, promises to remedy the distortions. **Martin Maiden's** subject is a set of Ibero-Romance 'empty affixes', unstressed morphemes with no discernible semantic content, and how they have been analyzed over the 20th century, particularly in the work of the scholar who identified them as a serious theoretical issue, the late Yakov Malkiel. **John Charles Smith** takes up another classic Romance linguistic problem, past participle agreement, and examines it in the perspective of markedness theory. His historiographically informed paper suggests among other things that a markedness account of French past participle agreement is anticipated in Michel Bréal's 1897 *Essai de sémantique*. The section concludes with a paper by **Brian D. Joseph** looking at the major dialect groupings within Romanian and how each might fit within a Balkan areal perspective, and drawing historical conclusions concerning the position of Romanian within Romance.

Part Seven, "Germanic, Caucasian, and Asian Linguistics", opens with three papers on Germanic languages. In the first, **David J. Holsinger and Joseph C. Salmons**, taking their cue from a remark by Bloomfield on the Neogrammarians, argue for "a complete analysis of the residues" of Old High German umlaut as an originally phonological phenomenon, contrary to some recent approaches that would treat it as morphologically conditioned. The paper by the late **Robert A. Hall, Jr.** is his valediction on the most controversial text of modern North Germanic historical linguistics, the runes of the Kensington Stone, the authenticity of which Hall maintained for half a century in the face of widespread skepticism. Here he marshalls new evidence by applying Halliday and Hasan's theory of textual cohesion to the runic inscription. **Matsuji Tajima**, Koerner's sometime collaborator, examines the compound gerund in Early Modern English, making use not only of extensive textual evidence but of the theoretical views of

the Prague linguist Bohumil Trnka and his pre-structuralist predecessors G. O. Curme and Otto Jespersen, the last of whom has been the subject of recent work by Koerner. **Bernard Comrie** frames his study of gender affixes in the North-east Caucasian language Tsez (also known by the Georgian name Dido) within the question of the roles of synchronic and diachronic explanation in linguistic analysis. In this case, he finds that the relationship between gender-based noun classes and the phonological segments which partly correlate with them suggests an intricate interplay of synchronic and diachronic factors. The final two papers deal with widely separated questions in Asian linguistics. **Alexander Vovin** takes up the controversy over the reading of a particular Old Korean phonogram which differs drastically if read on the basis of Old Chinese rather than of Middle Chinese evidence. Vovin offers extensive, detailed evidence in favor of the character being a leftover of Old Chinese within a system that used mostly Middle Chinese readings. Lastly, **Paul Sidwell**, co-editor of a recent volume in Koerner's honor (Kirk & Sidwell 1998), reviews the phonological systems reconstructed for proto-languages in the Bahnaric family of Southeast Asia, a branch of Austroasiatic. The paper argues that a historiographical approach is necessary in work of this kind that compares not so much primary linguistic data as the systems constructed by linguists, which, in Sidwell's view, are inevitably historical products that reflect the context and circumstances in which they were written. An epigram from Kirk & Sidwell (1998) identifies the work of Konrad Koerner as the direct source of their view that "no phenomenon in nature, including in languages, can be understood if taken by itself, isolated from surrounding phenomena; ...any phenomenon *can be understood and explained* if considered in its inseparable connection with surrounding phenomena".

We hope to have provided a range of solid and interesting papers giving a sense of where the historiography of the language sciences and historical linguistics stand a quarter century after Konrad Koerner devoted his life to helping establish the former and revitalize the latter. The publisher, editors, contributors, and subscribers to the *tabula gratulatoria* join together in a rousing chorus of admiration, appreciation, and affection for one of the giants of modern linguistics as he completes his sixtieth year, and wish him decades more of intellectual, professional, and personal fulfilment to come.

Authors' addresses

Sheila Embleton
 Department of Languages, Literatures & Linguistics
 York University
 4700 Keele Street
 Toronto, Ontario, CANADA M3J 1P3
 embleton@yorku.ca

John E. Joseph
 Department of Theoretical and Applied Linguistics
 University of Edinburgh
 Edinburgh EH8 9LN, United Kingdom
 John.Joseph@ed.ac.uk

Hans-Josef Niederehe
 Fachbereich II (Romanistik)
 Universität Trier
 D-54286 Trier
 niedereh@uni-trier.de

References

Note: For references to work by E. F. K. Koerner, see the Koerner Bibliography which follows.

- Aarsleff, Hans. 1970. "The History of Linguistics and Professor Chomsky". *Language* 46. 570–585.
- . 1982. *From Locke to Saussure: Essays on the study of language and intellectual history*. Minneapolis: University of Minnesota Press.
- Chomsky, Noam. 1966. *Cartesian Linguistics: A chapter in the history of rationalist thought*. New York: Harper & Row.
- Collinder, B. 1962. "Les origines du structuralisme". *Acta Societatis Linguisticae Uppsaliensis* 1. 1–15.
- Coseriu, Eugenio. 1967. "Georg von der Gabelentz et la linguistique synchronique". *Word* 23. 74–100. (Slightly modified version in Gabelentz 1969 [1901], 5–40.)
- Cowan, William & Michael K. Foster (eds.) 1989. *E. F. Konrad Koerner Bibliography*. Bloomington, Indiana: Eurolingua.
- Culler, Jonathan. 1975. *Saussure*. Hassocks: Harvester Press.

- Doroszewski, Witold. 1933a. "Sociologie et linguistique: Durkheim et de Saussure". *Actes du Deuxième Congrès International de Linguistes (Genève, 25-29 août 1931)*, 146-147. Paris: A. Maisonneuve.
- . 1933b. "Quelques remarques sur les rapports de la sociologie et la linguistique: Durkheim et F. de Saussure". *Journal de Psychologie* 30.82-91. Repr. in *Essais sur le langage*, ed. by Jean-Claude Pariente, 99-109, Paris: Minuit, 1969.
- Gabelentz, Georg von der. 1901. *Die Sprachwissenschaft: Ihre Aufgaben, Methoden und bisherigen Ergebnisse*. Repr. Tübingen: Tübinger Beiträge zur Linguistik, 1969 (including modified version of Coseriu 1967).
- Jakobson, Roman & Krystyna Pomorska. 1980. *Dialogues*. Paris: Flammarion. Engl. transl. by Christian Hubert, Cambridge, Mass.: MIT Press, 1983.
- Joseph, John E. 1988. "Saussure's Meeting with Whitney, Berlin, 1879". *Cahiers Ferdinand de Saussure* 42.205-214.
- . 1989. Review of Koerner, *Saussurean Studies/Études saussuriennes* (Geneva: Slatkine, 1988), *Language* 65.595-602.
- . 1991. Review of Koerner, *Practicing Linguistic Historiography: Selected essays* (Amsterdam & Philadelphia: John Benjamins, 1989), *Word* 42.216-219.
- . 1995. "The Structure of Linguistic Revolutions". *Historiographia Linguistica* 22.379-399.
- . 1996a. "'Undoubtedly a Powerful Influence': Victor Henry's *Antinomies linguistiques* (1896), with an annotated translation of the first chapter". *Language and Communication* 16.117-44.
- . 1996b. "The Immediate Sources of the 'Sapir-Whorf Hypothesis'". *Historiographia Linguistica* 23.365-404.
- Kirk, Neile & Paul Sidwell eds. 1998. *Practicing Koernerian Linguistics: A selection of papers and reviews presented in honour of Professor E. F. K. Koerner*. Melbourne: Association for the History of Language.
- Koerner, E. F. Konrad — see Koerner Bibliography (following section).
- Lieb, Hans-Heinrich, ed. 1992. *Prospects for a New Structuralism*. Amsterdam & Philadelphia: John Benjamins.
- Paul, Hermann. 1880. *Principien der Sprachgeschichte*. Halle/S.: Max Niemeyer. Engl. transl., *Principles of the History of Language*, by H. A. Strong, London: Swan Sonnenschein, Lowrey & Co., 1888.
- Saussure, Ferdinand de (1916) *Cours de linguistique générale*, publié par Charles Bally & Albert Sechehaye avec la collaboration d'Albert Riedlinger. Paris & Lausanne: Payot. 2nd, ed., 1922. Critical, ed. by Tullio De Mauro, Paris: Payot, 1973. English transl., *Course in General Linguistics*, by Roy Harris, London: Duckworth, 1983.

Bibliography of writings by E. F. K. Koerner, 1968–1999

Editors' note. This listing omits the volumes of *Historiographia Linguistica* (abbreviated below as *HL*) edited by Koerner from 1974 to the present and those of *Diachronica* edited or co-edited by him from 1984 to the present, as well as his many hundreds of brief book notices and editorial contributions within those volumes. Bibliographies and indices he has often painstakingly compiled for edited works are not mentioned, on the grounds that they are part of an editor's job, though few editors have ever been so thorough.

A particular difficulty has been faced with regard to articles on favorite topics revisited over the years. The practice we have followed is to list such articles separately, with cross-referencing, unless they have appeared in the same year in more or less the same form, in which case they are listed as a single item. Items of the same category published within a single book or journal issue are listed as a single item.

For the years 1968–88 we owe a great debt to the *E. F. Konrad Koerner Bibliography* edited by William Cowan & Michael K. Foster (Bloomington, Indiana: Eurolingua, 1989). We have not however adopted the numbering system used there, which corresponds to Koerner's own system for listing his works. The reason is that it lists them not in order of publication, as is the normal practice, but (apparently) in order of composition. Cowan & Foster do give publication dates, but whenever there is a discrepancy between the date printed on the work and the time it actually appeared, more importance is given to the latter, leading to confusion for the user. For example, taking at random p. 24 of Cowan & Foster, we find there seven entries (numbered 2.37 through 2.43) with the following publication dates: 2.37 (1979), 2.38 (1981), 2.39 (1979), 2.40 (1980), 2.41 ('1982 for 1981' [i.e., dated 1981 but appeared in 1982]), 2.42 (1984), 2.43 (1982 for 1981). In our bibliography they appear in the order 1979b (= 2.37), 1979c (= 2.39), 1980c (= 2.40), 1981a (= 2.38), 1981b (= 2.41), 1981c (= 2.43), 1984b (2.42). Clearly, the Cowan & Foster system will be of great use to any serious student

wishing to trace the development of Koerner's thinking from paper to paper. By the same token, the bibliography we have prepared makes it easier to trace the history of his work from the point of view of its public reception. In our listing, within each category of publications for a given year we have retained the relative ordering in Cowan & Foster unless there was some compelling reason to change it.

Abbreviations for book series directed by Koerner (all published by John Benjamins):

ACiL = Amsterdam Classics in Linguistics, 1800–1925

CiPL = Classics in Psycholinguistics

CILT = Current Issues in Linguistic Theory

LISL = Library and Information Sources in Linguistics

SiHoLS = Studies in the History of the Language Sciences

1968

Der Konjunktivgebrauch: Untersuchungen zum Modalsystem der deutschen Gegenwartssprache. [The Use of the Subjunctive: Investigations into the modal system of contemporary German]. M.A. thesis, Justus Liebig Universität Giessen, 1968. Unpublished. 2 parts, 148 pp.

1969

a) Rev. of Herwig Krenn & Klaus Müllner, *Bibliographie zur Transformationsgrammatik* (Heidelberg: Carl Winter, 1968). *Kratylos* 14:1.7–12.

b) Rev. of Georges Mounin, *Saussure, ou le structuraliste sans le savoir* (Paris: Seghers, 1968). *Canadian Journal of Linguistics/Revue canadienne de linguistique* 15:1.27–32.

1970

a) "Bloomfieldian Linguistics and the Problem of 'Meaning': A chapter in the history and study of language". *Jahrbuch für Amerikastudien/German Yearbook of American Studies* 15.162–183 (Heidelberg). Repr. in (1978a: 155–176) and in

Appraisals of Leonard Bloomfield, ed. by John G. Fought (London & New York: Routledge, 1999).

b) "Franz Nikolaus Finck (1867–1910): Zur 60. Wiederkehr des Todestages eines grossen Sprachwissenschaftlers aus dem Niederrhein". *Der Niederrhein: Zeitschrift für Heimatpflege* 37:3.91–94 (Krefeld).

c) "Ferdinand de Saussure — Origin and Development of his Linguistic Theory and its Influence upon the Major Linguistic Schools in the Western World: A critical evaluation of the relevance of Saussurean principles to contemporary theories of language". *Linguistische Berichte* no.9, 52–54.

d) Rev. art. on Peter H. Salus ed., *On Language: Plato to von Humboldt* (New York: Holt, Rinehart & Winston, 1969). *Lingua* 25:4.419–431.

e) Rev. of André Martinet ed., *La Linguistique: Guide alphabétique* (Paris: Denoël, 1969). *Linguistics* no. 61, 103–105.

f) Rev. of Herwig Krenn & Klaus Müllner comps., *Bibliographie zur Transformationsgrammatik* (Heidelberg: Carl Winter, 1968). *Language* 46:1.125–126.

1971

a) *Ferdinand de Saussure: Origin and Development of his Linguistic Theory in Western Studies of Language: A critical evaluation of the evolution of Saussurean principles and their relevance to contemporary linguistic theories*. Ph.D. dissertation, Simon Fraser University, Burnaby/Vancouver, B.C., Canada. [Rev. version publ. as (1973a).]

b) "Notes on the Semantics of Technical Terms in the Description of the Varieties of Contemporary German". *German Quarterly* 44:1.1–23.

c) "A Note on Transformational-Generative Grammar and the Saussurean Dichotomy of Synchrony versus Diachrony". *Linguistische Berichte* 13. 25–32.

d) Rev. of Rudolf Engler, *Lexique de la terminologie saussurienne* (Utrecht & Antwerp: Spectrum, 1968). *Language* 47:2.447–450.

e) Rev. of Harrison T. Meserole et al., *1969 MLA International Bibliography of Books and Articles on Modern Languages and Literatures*, vol.3: *Linguistics* (University Park, Pennsylvania: Pennsylvania State University Press, 1970). *Language* 47:4.915–918.

f) Rev. of Georg von der Gabelentz, *Die Sprachwissenschaft: Ihre Aufgaben, Methoden und bisherigen Ergebnisse*, reprint of 1901 ed. with preface by Gunter Narr & Uwe Petersen and an article by Eugenio Coseriu (Tübingen: Tübinger Beiträge zur Linguistik, 1969). *Lingua* 28:1/2.153–159.

- g) Rev. of Eberhard Zwirner & Kurt Zwirner, *Principles of Phonometrics*, transl. by Herman Bluhme (University, Alabama: University of Alabama Press, 1970). *Phonetica* 24:4.247–252.
- h) Brief rev. of Georges Mounin, *Saussure: Presentación y textos*, transl. by Juan Argente (Barcelona: Anagrama, 1969). *Romance Philology* 25:2.254–255.
- i) Rev. notes on Michael Lane ed., *Structuralism: A reader* (London: Cape, 1970) and Jean-Claude Pariente ed., *Essais sur le langage* (Paris: Minuit, 1969). *Germanistik* 12:2.207–208, 211–112.

1972

- a) *Bibliographia Saussureana 1870–1970: An annotated, classified bibliography on the background, development, and actual relevance of Ferdinand de Saussure's general theory of language*. Metuchen, N.J.: Scarecrow Press. 406 pp.
- b) *Contribution au débat post-saussurien sur le signe linguistique: Introduction générale et bibliographie annotée*. (Approaches to Semiotics, paperback series, 2.) The Hague: Mouton. 103 pp.
- c) "Glossaries of Linguistic Terminology, 1951–1971: An overview". *Linguistische Berichte* 18.30–38.
- d) "Hermann Paul and Synchronic Linguistics". *Lingua* 29:3.274–307. Repr. in (1978a: 73–106) and (1988a: 17–50).
- e) "Towards a Historiography of Linguistics: 19th and 20th century paradigms". *Anthropological Linguistics* 14:7.255–280. [Rev. & expanded version publ. as (1976c).]
- f) "Jan Baudouin de Courtenay: His place in the history of linguistic science". *Canadian Slavonic Papers/Revue canadienne des Slavistes* 14:4.663–682. Repr. in (1978a: 107–126).
- g) "A Note on Collinder on Chomsky". *Linguistics* 94.37–42.
- h) "A Brief Reply to Messrs. Narr and Petersen". *Lingua* 30:4.462–463. [Reply to Gunter Narr and Uwe Petersen, "On Koerner, Coseriu, and Gabelentz (A Brief rejoinder)", written in response to (1971i) and published together with this counterreply by Koerner.]
- i) Rev. of Hans Arens, *Sprachwissenschaft: Der Gang ihrer Entwicklung von der Antike bis zur Gegenwart*, 2nd ed. (Freiburg & München: Karl Alber, 1969). *Language* 48:2.428–445. Repr. in (1978a: 3–20).
- j) Rev. of Giulio C. Lepschy, *Die strukturelle Sprachwissenschaft: Eine Einführung*, transl. and with an additional chapter, by Harro Stammerjohann (München: Nymphenburg, 1969). *Göttingische Gelehrte Anzeigen* 224:3/4.288–295.

- k) Rev. of Frank Häusler, *Das Problem Phonetik und Phonologie bei Baudouin de Courtenay und in seiner Nachfolge* (Leipzig: Niemeyer, 1968). *Linguistics* 77.63–77.
- l) Rev. of Giulio C. Lepschy, *A Survey of Structural Linguistics* (London: Faber, 1970). *Linguistics* 91.93–101.
- m) Rev. of Robert Godel ed., *A Geneva School Reader in Linguistics* (Bloomington & London: Indiana University Press, 1969). *Vox Romanica* 31:1.116–124.
- n) Rev. of Theodor Siebs, *Deutsche Aussprache: Reine und gemässigte Hochlautung, mit Aussprachewörterbuch*, 19th rev. ed. by Helmut de Boor et al. (Berlin: W. de Gruyter, 1969). *German Quarterly* 45:1/2.196–199.
- o) Rev. of R. H. Robins, *Diversions of Bloomsbury: Selected writings on linguistics* (Amsterdam: North-Holland, 1970). *International Journal of American Linguistics* 38:2.151–154.
- p) Rev. of Ferdinand de Saussure, *Cours de linguistique générale*, critical ed. by Rudolf Engler, tome I (Wiesbaden: Harrassowitz, 1968). *Language* 48:3.682–684.
- q) Rev. of Wolfgang Kühlwein ed., *Linguistics in Great Britain*, vol. 2: *Contemporary Linguistics* (Tübingen: Niemeyer, 1970). *Journal of Linguistics* 8:2.342–345.
- r) Rev. of John T. Waterman, *Perspectives in Linguistics: An account of the background of modern linguistics*, 2nd ed. (Chicago & London: University of Chicago Press, 1970). *General Linguistics* 12:2.138–142.
- s) Rev. of Rolf Hiersche, *Ferdinand de Saussures Langue-Parole-Konzeption und sein Verhältnis zu Durkheim und von der Gabelentz* (Innsbruck: Institut für Sprachwissenschaft, Universität Innsbruck, 1972). *Kratylos* 17:2.125–129.

1973

- a) *Ferdinand de Saussure: Origin and Development of His Linguistic Thought in Western Studies of Language. A contribution to the history and theory of linguistics.* (Schriften zur Linguistik, 7.) Braunschweig: Friedrich Vieweg & Sohn. xl + 428 pp. [Rev. version of (1971a).] Hungarian transl., "Ferdinand de Saussure", by László Antál, in *Nyelvelméleti szakszöveggyűjtemény* [Text collection in linguistic theory], III: *A genfi iskola* [The Geneva School], II:411–740, Budapest: Tankönyvkiadó, 1982. Japanese transl., *Saussure no Gengoron* [Saussure's linguistic thought], by Kei-ichi Yamanaka, Tokyo: Taishukan, 1982. Spanish transl., *Ferdinand de Saussure: Génesis y evolución de su pensamiento en el marco de la lingüística occidental. Contribución a la historia y la teoría de la lingüística*, by Graciela García Montaña (Biblioteca Románica Hispánica, III: Manuales, 54), Madrid: Gredos, 1982.

- b) *The Importance of Techmer's Internationale Zeitschrift für Allgemeine Sprachwissenschaft in the Development of General Linguistics*. (SiHoLS, 1.) Amsterdam: John Benjamins. vii + 76 pp. Also published (with a different preface) with the reprint of *Internationale Zeitschrift für Allgemeine Sprachwissenschaft*, Supplement I: A. F. Pott, *Zur Literatur der Sprachenkunde Europas*. Amsterdam: John Benjamins, 1973.
- c) Rev. of André Jacob ed., *100 Points de vue sur le langage: 270 textes choisis et présentés, avec introduction et bibliographie* (Paris: Klincksieck, 1969). *Linguistics* 95.86-94.
- d) Rev. of J. J. Beylsmit et al., *Bibliographie linguistique pour l'année 1967 / Linguistic Bibliography for the Year 1967* (Utrecht & Antwerp: Spectrum, 1969). *Linguistics* 102.94-100.
- e) Rev. of Ferdinand de Saussure, *Corso di linguistica generale*, ed. by Tullio De Mauro, 3rd. ed. (Bari: Laterza, 1970). *Romance Philology* 27:1.75-78.
- f) Rev. of Maurice Leroy, *Les grands courants de la linguistique moderne*, 2nd ed. (Brussels: Éds. de l'Université de Bruxelles, 1971). *General Linguistics* 13: 1.54-56.
- g) Rev. of Wilhelm von Humboldt, *Linguistic Variability and Intellectual Development*, transl. by George C. Buck & Frithjof A. Raven (Coral Gables, Florida: University of Miami Press, 1971). *Language* 49:3.682-685.
- h) Rev. of Edward Stankiewicz ed. & transl., *A Baudouin de Courtenay Anthology* (Bloomington & London: Indiana University Press, 1972). *Language Sciences* 27.45-50.
- i) Rev. of Peter H. Salus, *Panini to Postal: A bibliography in the history of linguistics* (Edmonton, Alberta & Champaign, Illinois: Linguistic Research, 1971). *Foundations of Language* 10:4.589-594.
- j) Rev. of Paul Miclaŭ, *Le signe linguistique* (Bucharest: Éds. de l'Académie de la RSR; Paris: Klincksieck, 1970). *Linguistics* 117.123-125.
- k) Rev. of Michael Silverstein ed., *Whitney on Language: Selected writings of William Dwight Whitney*, with an introductory essay by Roman Jakobson (Cambridge, Mass.: MIT Press, 1971). *Journal of the American Oriental Society* 93:4.617-619.

1974

- a) (Editor), Franz Bopp: *Analytical Comparison of the Sanskrit, Greek, Latin, and Teutonic Languages, shewing the original identity of their grammatical structure*. With a preface by the editor. (ACiL, 3.) Amsterdam: John Benjamins. xxxviii +

- 68 pp. (2nd corrected ed. with new foreword by the editor, Amsterdam & Philadelphia: John Benjamins, 1989.)
- b) (Editor), August Friedrich Pott: *Einleitung in die Allgemeine Sprachwissenschaft und Zur Literatur der Sprachenkunde Europas*. With a preface by the editor. (ACiL, 10.) Amsterdam: John Benjamins. xlv + 462 pp.
- c) (Editor), Berthold Delbrück: *Introduction to the Study of Language: A critical survey of the history and methods of comparative philology of Indo-European languages*. With a foreword by the editor. (ACiL, 8.) Amsterdam: John Benjamins. xix + 148 pp. (2nd printing, with corrections and an editorial note, 1989.)
- d) "Paradigms in the 19th and 20th Century History of Linguistics: Schleicher — Saussure — Chomsky". *Proceedings of the Eleventh International Congress of Linguists*, ed. by Luigi Heilmann, I:121–132. Bologna: Il Mulino.
- e) "Animadversions on Some Recent Claims regarding the Relationship between Georg von der Gabelentz and Ferdinand de Saussure". *Studi saussuriani per Robert Godel*, ed. by René Amacker, Tullio De Mauro et al., 165–180. Bologna: Il Mulino. Repr. in (1978a: 137–152) and in (1988a: 51–66).
- f) "Purpose and Scope of 'Historiographia Linguistica': Editorial". *HL* 1:1.1–10. Partly repr. in (1978a: 55–62).
- g) "An Annotated Chronological Bibliography of Western Histories of Linguistic Thought, 1822–1972. Part I: 1822–1915, Part II: 1916–1961, Part III: 1962–1972". *HL* 1:1.81–94, 1:2.185–205, 1:3.351–384. [See also (1975c).] Rev. versions publ. in (1978b).
- h) Rev. of Johann Severin Vater, *Literatur der Grammatiken, Lexica und Wörter-sammlungen aller Sprachen der Erde*, repr. of 1847 ed. by Bernhard Jülg (Graz: Akad. Druck- u. Verlagsanstalt, 1970). *Linguistics* 131.118–121.
- i) Rev. of Wolfgang Kühlwein ed., *Linguistics in Great Britain*, vol. I: *History of Linguistics* (Tübingen: Niemeyer, 1971). *Archiv für das Studium der neueren Sprachen und Literaturen* 211:2.108–115.
- j) "Carl Svedelius (1861–1951)", Obituaries of Waltraud Bumann (1936–1973) and Louis Kukenheim Ezn (1905–1972). *HL* 1:1.139–140, 283–284, 285–286.
- k) (Editor), Henri Gaidoz: "Carl Lottner (1834–73)". *HL* 1:2.287–290.
- l) (Editor), Henri Cordier: "Georg von der Gabelentz (1840–1893)". *HL* 1:3.427–431.
- m) "Preface" to repr. of Wilhelm Bacher, *Die Anfänge der hebräischen Grammatik und Die hebräische Sprachwissenschaft vom 10. bis zum 16. Jahrhundert*, v–vii. (SiHoLS, 4). Amsterdam: John Benjamins.

1975

- a) (Editor, with the assistance of John Odmark & J. Howard Shaw), *The Transformational-Generative Paradigm and Modern Linguistic Theory*. (CILT, 1.) Amsterdam: John Benjamins. viii + 462 pp.
- b) "European Structuralism: Early beginnings". *Current Trends in Linguistics* ed. by Thomas A. Sebeok, vol. 13: *Historiography of Linguistics*, II:717-827. The Hague: Mouton.
- c) "An Annotated Chronological Bibliography of Western Histories of Linguistic Thought, 1822-1972. Part IV: Addenda". *HL* 2:2.207-222 (1975). [See (1974g).] Rev. version publ. in (1978b).
- d) Rev. of Iorgu Iordan, *An Introduction to Romance Linguistics, its Schools and Scholars*, transl. and rev. by John Orr; 2nd ed., with a supplement by Rebecca Posner, "Thirty Years on" (Oxford: Blackwell, 1970). *Linguistics* 143.103-107.
- e) Rev. notes on Hans Helmut Christmann, *Idealistische Philologie und moderne Sprachwissenschaft* (München: Wilhelm Fink, 1974); Rudolf Engler ed., *Ferdinand de Saussure: Cours de linguistique générale*, fascicles 3+4 (Wiesbaden: Harrassowitz, 1968-74); and Hans-Dieter Kreuder, *Studienbibliographie Linguistik* (Wiesbaden: Steiner, 1974). *Germanistik* 16:2.320-321, 357, 345.
- f) Rev. notes on Joann Vogt, *The Linguistic Work of Friedrich Karl Fulda* (The Hague: Mouton, 1974) and Tasso Borbé ed., *Zur Kritik der marxistischen Sprachtheorie N. Ja. Marr's* (Kronberg/Taunus: Scriptor, 1974). *Germanistik* 16:3.580, 583-584.
- g) Rev. note on Gudula Dinser ed., *Zur Theorie der Sprachveränderung* (Kronberg/Taunus: Scriptor, 1974) and Federico Albano Leoni, *Il primo Trattato Grammaticale Islandese* (Bologna: Il Mulino, 1975). *Germanistik* 16:4.773, 781.
- h) "Philipp Aronstein (4 Dec. 1862-24 Sept. 1942)" and "The Concept of Structure in General Linguistic Theory; its evolution from the beginnings of the New Philology to contemporary linguistic doctrines: A contribution to the epistemology and historiography of linguistic science". *HL* 2:1.126-129, 131-134.

1976

- a) (Editor), Victor Hehn: *Cultivated Plants and Domesticated Animals in their Migration from Asia to Europe: Historico-linguistic studies*. With a foreword by the editor. (ACiL, 7.) Amsterdam: John Benjamins. lxxv + 523 pp.

- b) "Zu Ursprung und Geschichte der Besternung in der historischen Sprachwissenschaft: Eine historiographische Notiz". *Zeitschrift für vergleichende Sprachforschung* 89:2.185–190. [Abridged version publ. as (1977c).] Repr. in (1978a: 211–216).
- c) "Towards a Historiography of Linguistics: 19th and 20th century paradigms". *History of Linguistic Thought and Contemporary Linguistics* ed. by Herman Parret, 685–718. Berlin & New York: W. de Gruyter. [Rev. & expanded version of (1972e).] Repr. in (1978a: 21–54).
- d) "The Importance of Linguistic Historiography and the Place of History in Linguistic Science". *Foundations of Language* 14:4.541–547. Repr. in (1978a: 63–69).
- e) "Saussure and the French Linguistic Tradition". In *Memoriam Friedrich Diez: Akten des Kolloquiums zur Wissenschaftsgeschichte der Romanistik* (Trier, 2.-4. Okt. 1975) ed. by Hans-J. Niederehe et al., 405–417. Amsterdam: John Benjamins.
- f) "A Minor Figure in 19th-Century French Linguistics: A. Dufriche-Desgenettes". *Phonetica* 33:3.222–231. Repr. in (1978a: 127–136).
- g) "1876 as a Turning Point in the History of Linguistics". *Journal of Indo-European Studies* 4:4.333–353. Repr. in (1978a: 189–209).
- h) Rev. note on Valentin N. Voloshinov, *Marxismus und Sprachphilosophie*, transl. by Renate Horlemann and ed. with an intro. by Samuel M. Weber (Frankfurt/M.-Berlin-Wien: Ullstein, 1975). *Germanistik* 17:2.395–396.
- i) "Foreword" to repr. of Rasmus Kristian Rask, *A Grammar of the Icelandic or Old Norse Tongue* ed. by Thomas L. Markey, vii–viii. (ACiL, 2.) Amsterdam: John Benjamins.

1977

- a) (Editor), Friedrich Schlegel: *Über die Sprache und Weisheit der Indier: Ein Beitrag zur Begründung der Altertumskunde*. With a foreword by the editor. (ACiL, 1.) Amsterdam: John Benjamins. lvii + 190 pp.
- b) (Editor) Adolf Holtzmann: 'Ueber den Umlaut: Zwei Abhandlungen' and 'Ueber den Ablaut'. With a foreword by the editor. (ACiL, 12.) Amsterdam: John Benjamins. xxix, 48 + 81 pp.
- c) "Sulla origine e storia dell'asterisco: Un cenno storiografico". *Atti del Nono Convegno Internazionale di Studi sulla Ricostruzione in Linguistica* (Pavia, 1–2 Oct. 1975) ed. by Raffaele Simone & Ugo Vignuzzi, 253–258. Rome: Bulzoni. [Abridged version of (1976b).] Repr. with a postscript in (1989a: 179–184).

- d) "On the Non-Applicability of Kuhn's Paradigms to the History of Linguistics". *Proceedings of the Seventh Annual Meeting of the North Eastern Linguistic Society* (Cambridge, Mass., 5-7 November 1976) ed. by Judy Kegl, David Nash & Annie Zaenen, 165-174. Cambridge, Mass.: Department of Linguistics & Philosophy, MIT.
- e) "The Humboldtian Trend in Linguistics". *Studies in Descriptive and Historical Linguistics: Festschrift for Winfred P. Lehmann* ed. by Paul J. Hopper, 145-158. Amsterdam: John Benjamins. Also in *Cahiers Linguistiques d'Ottawa* 5.27-40 (1977).
- f) Rev. of Friedrich Schlegel, *Studien zur Philosophie und Theologie* (Paderborn: F. Schöningh, 1975). *Indogermanische Forschungen* 82.229-233.
- g) Rev. notes on Gordon Winant Hewes, *Language Origins: A bibliography* (The Hague: Mouton, 1975) and Peter Schifko, *Bedeutungstheorie: Einführung in die linguistische Semantik* (Stuttgart-Bad Cannstatt: Frommann-Holzboog, 1975) *Germanistik* 18:1.28-29, 38.
- h) Rev. note on Jonathan Culler, *Saussure* (London: Fontana/Collins, 1976). *Germanistik* 18:2.292.
- i) Rev. notes on Philip Pettit, *The Concept of Structuralism: A critical analysis* (Dublin: Gill & Macmillan, 1975); Frank Häusler, *Das Problem Phonetik und Phonologie bei Baudouin de Courtenay und in seiner Nachfolge*, 2nd ed. (Halle/Saale: Max Niemeyer, 1976); and G. A. Padley, *Grammatical Theory in Western Europe 1500-1700: The Latin tradition* (Cambridge: Cambridge University Press, 1976). *Germanistik* 18.3.618-691, 643-644, 653-654.
- j) Rev. note on Ian Robinson, *The New Grammarians' Funeral: A critique of Noam Chomsky's linguistics* (Cambridge: Cambridge University Press, 1975). *Germanistik* 18:4.927.
- k) "Kratzenstein, Christian Gottlieb (1723-1795)", "Saussure, Ferdinand de (1857-1913)", "Techmer, Friedrich (1843-1891)". *Biographical Dictionary of the Phonetic Sciences* ed. by Arthur J. Bronstein et al., 120-121, 183-185, 206-207. Bronx, N.Y.: Press of Lehman College.
- l) "Foreword" to *Analogy: A basic bibliography* by Raimo Anttila & Warren A. Brewer, v-viii. Amsterdam: John Benjamins, 1977.

1978

- a) *Toward a Historiography of Linguistics: Selected essays*. Foreword by R. H. Robins. (SiHoLS, 19.) Amsterdam: John Benjamins, xx + 222 pp. [Includes (1970a), (1972d, f & i), (1974e & f), (1976b, c, d, f & g), (1978c).]

- b) *Western Histories of Linguistics, 1822–1976: An annotated, chronological bibliography*. (SiHoLS, 11.) Amsterdam: John Benjamins. ix + 113 pp. (Includes rev. versions of (1974g) and (1975c), plus a section on 1973–76 publications.)
- c) “Zu Ursprung und Entwicklung des Phonem-Begriffs: Eine historische Notiz”. *Sprache in Gegenwart und Geschichte: Festschrift für Heinrich Matthias Heinrichs* ed. by Dieter Hartmann, Hansjürgen Linke & Otto Ludwig, 82–93. Vienna & Cologne: H. Böhlau. Repr. in (1978a: 177–188).
- d) “Toward a Historiography of Phonetics: A brief survey of the state of the art”. *Festschrift für Otto von Essen zu seinem 80. Geburtstag am 28. Mai 1978* ed. by Hans-Heinrich Wängler, 221–237. Hamburg: Helmut Buske. [Rev. & expanded version of (1979b). See also (1991d), (1993e).]
- e) “Aims and Methods of the History of Linguistics: Contrasting views”. *Proceedings of the Twelfth International Congress of Linguists (Vienna, Aug. 28–Sep. 2, 1977)* ed. by Wolfgang U. Dressler & Wolfgang Meid, 108–112. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.
- f) (with Matsuji Tajima), “Saussure in Japan: A survey of research, 1928–1978”. *HL* 5:1/2.121–148. Repr. in (1988a: 175–202).
- g) Rev. notes on Elisabeth Bense et al., eds. & transls., *Beschreibungsmethoden des amerikanischen Strukturalismus* (München: Hueber, 1976); Maria di Salvo, *Il pensiero linguistico di J. Baudouin de Courtenay* (Venice & Padua: Marsilio, 1975); and Linda R. Waugh, *Roman Jakobson's Science of Language* (Lissel Holland: P. de Ridder, 1976). *Germanistik*. 19:2.245–246, 250–251, 273.
- h) “Foreword” to repr. of *Experimentelle Untersuchungen über die psychologischen Grundlagen der sprachlichen Analogiebildungen* by Albert Thumb & Karl Marbe, ed. by David J. Murray, v–viii. (CiPL, 1.) Amsterdam: John Benjamins.
- i) “Foreword” to *The Theory of English Lexicography* by Tetsuro Hayashi (SiHoLS, 18.), v–viii. Amsterdam: John Benjamins.

1979

- a) “Pilot and Parasite Disciplines in the Development of Linguistic Science”. *The Fifth LACUS Forum 1978* ed. by Wolfgang Wölck, 525–534. Columbia, South Carolina: Hornbeam Press. [Also publ. as (1980c).]
- b) “Toward a Historiography of Phonetics”. *Current Issues in the Phonetic Sciences: Proceedings of the IPS-77 Congress, Miami Beach, Florida, 17–19 Dec. 1977* ed. by Harry & Patricia Hollien, I: 23–35. Amsterdam: John Benjamins. [Rev. & expanded version publ. as (1978d). See also (1991d), (1993e).]

c) "L'importance de William Dwight Whitney pour les jeunes linguistes de Leipzig et pour F. de Saussure". *Studies in Diachronic, Synchronic, and Typological Linguistics: Festschrift for Oswald Szemerényi on the occasion of his 65th birthday* ed. by Bela Brogyanyi, I:437–454. Amsterdam: John Benjamins. [Rev. version publ. as (1980e).]

1980

a) (Editor, with Hans-Josef Nedderhede & Robert Henry Robins), *Studies in Medieval Linguistic Thought, dedicated to Geoffrey L. Bursill-Hall on the occasion of his sixtieth birthday on 15 May 1980*. (SiHoLS, 26.) Amsterdam: John Benjamins. vi + 321 pp. [Editor's own contributions on pp. 7–11 and 265–299.]

b) (Editor), *Progress in Linguistic Historiography: Papers from the International Conference on the History of the Language Sciences (Ottawa, 28–31 August 1978)*. (SiHoLS, 20.) Amsterdam: John Benjamins. vi + 320 pp.

c) "Pilot and Parasite Disciplines in the Development of Linguistic Science". *Folia Linguistica Historica* 1:1.213–224. [Also publ. as (1979a).] Repr. in (1989a: 245–256), with 'Parasite' replaced by 'Pirate'.

d) "Sur l'origine du concept et du terme de 'synchronique' en linguistique". *Recherches de linguistique: Hommages à Maurice Leroy*, ed. by Jean Bingen, André Coupez & Francine Mawet, 100–109. Bruxelles: Éds. de l'Université de Bruxelles. Repr. in (1988a: 89–98) and (1989a: 257–266).

e) "L'importance de William Dwight Whitney pour les jeunes linguistes de Leipzig et pour Ferdinand de Saussure". *Linguisticae Investigationes* 4:2.379–394. [Rev. version of (1979c).] Repr. in (1988a: 1–16).

f) "Medieval Linguistic Thought: A comprehensive bibliography". *HL* 7:1/2.265–300.

1981

a) "The Neogrammarian Doctrine: Breakthrough or extension of the Schleicherian paradigm: A problem in linguistic historiography". *Folia Linguistica Historica* 2.2.157–178. [Rev. & expanded version of (1982c).] Repr. in (1989a: 79–100).

b) "Schleichers Einfluss auf Haeckel: Schlaglichter auf die wechselseitige Abhängigkeit zwischen linguistischen und biologischen Theorien im 19. Jahrhundert". *Nova Acta Leopoldina* N.F. 54, no. 245, 731–745. Leipzig: J. A. Barth.

Rev. version publ. in *Zeitschrift für vergleichende Sprachforschung* 95:1.1–21 (1981), and repr. in (1989a: 211–231).

c) “Observations on the Sources, Transmission, and Meaning of ‘Indo-European’ and Related Terms in the Development of Linguistics”. *Indogermanische Forschungen* 86.1–29. [Rev. & expanded version of (1982c).] Repr. in (1989a: 149–177).

d) Rev. of Hans-Helmut Christman ed., *Sprachwissenschaft des 19. Jahrhunderts* (Darmstadt: Wissenschaftliche Buchgesellschaft, 1977). *Indogermanische Forschungen* 86.292–300.

e) Rev. notes on Günter Dresselhaus, *Langue/Parole und Kompetenz/Performanz: Zur Klärung der Begriffspaare bei Saussure und Chomsky; ihre Vorgeschichte und ihre Bedeutung für die moderne Linguistik* (Frankfurt: Peter Lang, 1979) and Bernhard Gröschel, *Materialistische Sprachwissenschaft: Internationale analytische Bibliographie* (Weinheim & Basel: Beltz, 1978). *Germanistik* 22:1.21–22, 24–25.

f) Rev. note on Thomas M. Scheerer, *Ferdinand de Saussure: Rezeption und Kritik* (Darmstadt: Wissenschaftliche Buchgesellschaft, 1980). *Germanistik* 22:2.294–295.

1982

a) (Editor, with J. Peter Maher & Allan R. Bomhard), *Papers from the Third International Conference on Historical Linguistics* (Hamburg, 22–26 Aug. 1977). (CILT, 13.) Amsterdam: John Benjamins. xiii + 434 pp. [Editor’s own contributions on pp. 129–152 and 153–180.]

b) “Models in Linguistic Historiography”. *Forum Linguisticum* 6:3.189–201. Repr. in (1989a: 47–59).

c) “The Neogrammarian Doctrine: Breakthrough or Extension of the Schleicherian Paradigm: A problem in linguistic historiography” and “Observations on the Sources, Transmission, and Meaning of ‘Indo-European’ and Related Terms in the Development of Linguistics”. In (1982a), 129–152, 153–180. [Rev. & expanded versions publ. as (1981a & c)].

d) “Positivism in 19th-Century Linguistics”. *Rivista di Filosofia* 73.170–191.

e) “The Schleicherian Paradigm in Linguistics”. *General Linguistics* 22:1.1–39. [Rev. version publ. in (1983c: xxxii–lxii), further rev. & expanded version publ. in (1989a: 25–375).]

f) “On the Historical Roots of the Philology/Linguistics Controversy”. *Papers from the 5th International Conference on Historical Linguistics* ed. by Anders

- Ahlqvist, 404–413. Amsterdam: John Benjamins. Repr. (with new bibl.) in (1989a: 233–244). [Rev. version publ. as (1997b). See also (1982j).]
- g) "Positivism in Linguistics". *The Eighth LACUS Forum 1981* ed. by Waldemar Gutwinski, 82–99. Columbia, South Carolina: Hornbeam Press. Rev. & expanded version publ. in *Sprachwissenschaft* 7.359–377 (1982), and repr. in (1989a: 191–209).
- h) Preface to the Spanish translation of (1973a), 9–13.
- i) Preface to the Japanese translation of (1973a), iii–vi.
- j) Contribution to the "Discussion". *Papers from the 5th International Conference on Historical Linguistics* ed. by Anders Ahlqvist, 445–448. Amsterdam: John Benjamins. [See also (1982f).]
- k) "Comptes rendus des Colloques: XIIIth International Congress of Linguists (Tokyo, août 1982)". *Bulletin d'Information de la Société d'Histoire et d'Épistémologie des Sciences du Langage* no. 9, 12–18.

1983

- a) (Editor), Leonard Bloomfield, *An Introduction to the Study of Language*, with an intro. by Joseph F. Kess and a foreword by the editor. (CiPL, 3.) Amsterdam & Philadelphia: John Benjamins. xxxviii + x + 335 pp.
- b) (Editor, with Cornelis H. M. Versteegh & Hans-Josef Niederehe), *History of Linguistic Thought in the Near East*. Preface by K. Koerner. (SiHoLS, 28.) Amsterdam & Philadelphia: John Benjamins. xi + 265 pp.
- c) (Editor), August Schleicher, *Die Sprachen Europas in systematischer Übersicht: Linguistische Untersuchungen*. With preface and introductory article [rev. version of (1982e)] by the editor. (ACiL, 4.) Amsterdam & Philadelphia: John Benjamins. lxxi + x + 274 pp.
- d) (Editor), *Linguistics and Evolutionary Theory: Three Essays by August Schleicher, Ernst Haeckel, and Wilhelm Bleek*. With an intro. by J. Peter Maher and a foreword by the editor. (ACiL, 6.) Amsterdam & Philadelphia: John Benjamins. xlv + 82 + 78 pp.
- e) (Editor), Holger Pedersen, *A Glance at the History of Linguistics, with particular regard to the historical study of phonology*. Transl. by Caroline C. Henriksen, with an intro. by the editor. (SiHoLS, 7.) Amsterdam & Philadelphia: John Benjamins. xxxiii + 100 pp. [See also (1985e).]
- f) "Quelques observations au sujet des sources de la sémiologie saussurienne". *History of Semiotics* ed. by Achim Eschbach & Jürgen Trabant, 323–339. Amsterdam & Philadelphia: John Benjamins. [Rev. version publ. as (1985a).]

- g) "A Critique of Recent Histories of Linguistics". *Proceedings of the XIIIth International Congress of Linguists, August 29–September 1982, Tokyo* ed. by Shirô Hattori & Kazuko Inoue, 877–880. Tokyo. [See also (1983k).]
- h) "Saussure's French Connection". *Papers from the 12th Linguistic Symposium on Romance Languages* ed. by Philip H. Baldi, 475–500. Amsterdam & Philadelphia: John Benjamins. [Rev. version publ. as (1984d).]
- i) "The Chomskyan 'Revolution' and Its Historiography: A few critical comments". *Language & Communication* 3:2.147–169 (Oxford, 1983). [Rev. & expanded version of (1984c).]
- j) Rev. notes on Klaus Grotzsch, *Sprachwissenschaftsgeschichtsschreibung: Ein Beitrag zur Kritik und methodologischen Selbstvergewisserung der Disziplin* (Göppingen: Werner Kümmerle, 1982) and Hans Aarsleff, *From Locke to Saussure: Essays on the study of language and intellectual history* (London: Athlone, 1982). *Germanistik* 24:3/4.531–532, 540.
- k) "History (and Philosophy) of Science and Linguistics: Introductory remarks". *Proceedings of the XIIIth International Congress of Linguists, August 29–September 1982, Tokyo* ed. by Shirô Hattori & Kazuko Inoue, 1213–14. Tokyo. [See also (1982k), (1983g).]

1984

- a) (Editor), *Edward Sapir: Appraisals of his life and work*. With an intro. by the editor. (SiHoLS, 36.) Amsterdam & Philadelphia: John Benjamins. xxviii + 224 pp.
- b) "Franz Bopp, 14.9.1791–23.10.1867". *Aschaffener Jahrbuch für Geschichte, Landeskunde und Kunst des Untermaingebietes* 8.311/313–320. Aschaffenburg: Geschichts- und Kunstverein. Repr. with new bibl. in (1989a: 291–298).
- c) "The 'Chomskyan Revolution' and its Historiography". *Sincronia, diacronia e cultura: Saggi linguistici in onore di Luigi Heilmann* ed. by Enrico Arcaini et al., 153–177. Brescia: Editrice La Scuola. [Rev. & expanded version publ. as (1983i).]
- d) "French Influences on Saussure". *Canadian Journal of Linguistics/Revue Canadienne de Linguistique* 29:1.20–41. [Rev. version of (1983h).] Repr. in (1988a: 67–68).
- e) "Karl Bühler's Theory of Language and Ferdinand de Saussure's 'Cours'". *Lingua* 62.3–24. Repr. in (1988a: 115–136). German transl., "Karl Bühler's Sprachtheorie und Ferdinand de Saussure's *Cours*", by Achim Eschbach, in *Bühler-Studien* ed. by Achim Eschbach, II:89–115. Frankfurt/Main: Suhrkamp, 1984.

- f) "Kruszewski's Contribution to General Linguistic Theory". *The Tenth LACUS Forum 1983* ed. by Alan Manning, Pierre Martin & Kim McCalla, 65–78. Columbia, South Carolina: Hornbeam Press. [Rev. version published as (1986d).]
- g) "Remarques critiques sur la linguistique américaine et son historiographie". *Linguisticae Investigationes* 8:1.87–103.
- h) "Report on the Sixth International Conference on Historical Linguistics (Poznań, 22–26 Aug. 1984)". *Diachronica* 1.145–150.

1985

- a) "Quelques observations sur les sources de la sémiologie saussurienne". *Lingua e Stile* 20:3.287–301. [Rev. version of (1983f).] Repr. in (1988a: 99–113).
- b) "The Place of Saussure's *Mémoire* in the Development of Historical Linguistics". *Papers from the Sixth International Conference on Historical Linguistics* ed. by Jacek Fisiak, 323–345. Amsterdam & Philadelphia: John Benjamins; Poznań: Adam Mickiewicz University Press. [Rev. version publ. as (1987c).]
- c) "On Writing the History of American Linguistics". *The Eleventh LACUS Forum 1984* ed. by Robert A. Hall, Jr., 531–542. Columbia, South Carolina: Hornbeam Press.
- d) "Aperçu historique de l'aspect social en linguistique". *Bulletin de l'Association Canadienne de Linguistique Appliquée* 7:2.113–127. [Rev. version publ. as (1986e). See also (1986f).]
- e) "Holger Pedersen: A sketch of his life and work". *Critica Storica* 22.236–252. [Rev. version of intro. to (1983e).] Repr. in (1989a: 417–433).
- f) Rev. of Manfred Kohrt, *Phonetik, Phonologie und die "Relativität der Verhältnisse": Zur Stellung Jost Wintelers in der Wissenschaft* (Wiesbaden: Franz Steiner, 1984). *Word* 36:3.258–265. Repr. with the new title "On Practicing Linguistic Historiography" in (1989a).

1986

- a) (with Matsuji Tajima and the collaboration of Carlos P. Otero), *Noam Chomsky: A personal bibliography, 1951–1986*. (Library and Information Sources in Linguistics, 11.) Amsterdam & Philadelphia: John Benjamins. xi + 217 pp.
- b) (Editor, with William Cowan & Michael K. Foster), *New Perspectives in Language, Culture, and Personality: Proceedings of the Edward Sapir Centenary Conference (Ottawa, 1–3 Oct. 1984)*. (SiHoLS, 41.) Amsterdam & Philadelphia: John Benjamins. xiv + 627 pp.

- c) (Editor, with Hans-Josef Niederehe & Paolo Ramat), *History of Linguistics in Italy*. (SiHoLS, 33.) Amsterdam & Philadelphia: John Benjamins. x + 363 pp.
- d) "Kruszewski's Contribution to General Linguistic Theory". *Linguistics across Historical and Geographical Boundaries: In honour of Jacek Fisiak* ed. by Dieter Kastovsky & Aleksander Szwedek, vol. I: *Linguistic Theory and Historical Linguistics*, 53–75. Berlin & New York: Walter de Gruyter. [Rev. & expanded version of (1984f).] Repr. in (1989a: 377–399). Russian transl., "Vklad Kruševskogo v obščē-lingvističeskiju teoriju", by Fedor M. Berezin, in N. V. Krushevskij, *Izbrannye roboty po jazykoznaniju* [Collected works in linguistics] ed. by Fedor M. Berezin, Moscow: Nasledie, 1998.
- e) "Petit aperçu historique de l'aspect social en linguistique". *Cahiers Linguistiques d'Ottawa* 15.95–121. [Rev. version of (1985d). See also (1986f).]
- f) "Aux sources de la sociolinguistique". *Lingvisticae Investigationes* 10:2.381–401. [Rev. & expanded version of (1985d) and (1986e). Another version publ. as (1989c); see also (1989e), (1991c).] Repr. in (1988a: 155–174).
- g) Rev. notes on Arno Beyer, *Deutsche Einflüsse auf die englische Sprachwissenschaft im 19. Jahrhundert* (Göppingen: Werner Kümmerle, 1981) and Peter Schmitter, *Untersuchungen zur Historiographie der Linguistik: Struktur — Methodik — theoretische Fundierung* (Tübingen: Gunter Narr, 1982). *Germanistik* 27:2.226–227, 241–242.

1987

- a) (Co-editor, with Robert A. Hall, Jr., as main editor), *Leonard Bloomfield: Essays on his life and work*. (SiHoLS, 47.) Amsterdam & Philadelphia: John Benjamins. x + 237 pp.
- b) "On Schleicher and Trees". *Biological Metaphor and Cladistic Classification: An interdisciplinary approach* ed. by Henry M. Hoenigswald & Linda F. Wiener, 109–113. Philadelphia: University of Pennsylvania Press. Repr. in (1989a: 185–189) with the title "August Schleicher and the Tree Idea in Comparative Linguistics".
- c) "The Place of Saussure's *Mémoire* in the Development of Historical Linguistics". *Festschrift for Henry M. Hoenigswald* ed. by George Cardona & Norman Zide, 201–217. Tübingen: Gunter Narr. [Rev. version of (1985b).] Repr. in (1988a: 137–153).
- d) "Das Problem der Metasprache in der Sprachwissenschaftsgeschichtsschreibung". *Zur Theorie und Methode der Geschichtsschreibung der Linguistik: Analysen und Reflexionen* ed. by Peter Schmitter, 63–80. Tübingen: Gunter Narr. Repr. in (1989a: 13–30).

- e) "On the Problem of 'Influence' in Linguistic Historiography". *Papers in the History of Linguistics: Proceedings of the Third International Conference on the History of the Language Sciences (ICHoLS III), Princeton, 19-23 Aug. 1984* ed. by Hans Aarsleff, Louis G. Kelly & Hans-Josef Niedereche, 13-28. Amsterdam & Philadelphia: John Benjamins, 1987. Repr. in (1989a: 31-46). Portuguese transl., "O problema da 'influência' em historiografia lingüística", by Cristina Altman, in *Investigações: Lingüística e Teoria Literária. Revista do Programa de Pós-Graduação em Letras e Lingüística da University Federal de Pernambuco* 8.181-213 (Recife, Brazil, 1998).
- f) "Friedrich Schlegel and the Emergence of Historical-Comparative Grammar". *Lingua e Stile* 22:3.341-365. [Rev. & expanded version of (1990f). Further expanded version publ. in (1989a: 269-290).]

1988

- a) *Saussurean Studies/Études saussuriennes*. Avant-propos de Rudolf Engler. Geneva: Éditions Slatkine. xxii + 207 pp. [Includes (1972d), (1974e), (1978f), (1980d & e), (1984d & e), (1985a), (1986f), (1987c).]
- b) "Jacob Grimm's Place in the Foundation of Linguistics as a Science". *Word* 39:1.1-20. [Rev. version of (1990e). Further expanded version publ. in (1989a: 303-323).]
- c) "Toward a History of Americanist Linguistics". *The Fourteenth LACUS Forum 1987* ed. by Sheila Embleton, 336-343. Lake Bluff, Illinois: LACUS. Also publ. as (1989f). [Rev. version publ. as (1992d). See also (1994d).]
- d) "Meillet, Saussure et la linguistique générale". *Antoine Meillet et la linguistique de son temps* ed. by Sylvain Auroux (= *Histoire-Épistémologie-Langage* 10:2), 57-73. Paris: Société d'Histoire et d'Épistémologie des Sciences du Langage. [Rev. & expanded version publ. in (1989a: 401-415).]
- e) Rev. note on Roy Harris, *Reading Saussure: A critical commentary on the Cours de linguistique générale* (London: Gerald Duckworth, 1987). *Germanistik* 29.310-311.

1989

- a) *Practicing Linguistic Historiography: Selected essays*. (SiHoLS, 50.) Amsterdam & Philadelphia: John Benjamins. xii + 455 pp. [Includes (1977c), (1980c, d), (1981a, b & c), (1982b, e, f & g), (1984b), (1985e & f), (1986d), (1987b, d, e &

- f), (1988b & d), (1989d), (1990e, f & g); several of these have been substantially recast and expanded].
- b) "Comments on Reconstructions in Historical Linguistics". *The New Sound of Indo-European: Essays in phonological reconstruction* ed. by Theo Vennemann, 3–15. Berlin & New York: Walter de Gruyter.
- c) "Aux sources de la sociolinguistique". *Actes du XVIIIe Congrès International de Linguistique et de Philologie Romanes: Université de Trèves (Trier) 1986*, ed. by Dieter Kremer, tome VII: *Histoire de la linguistique et de philologie romanes*, 16–33; discussion, 33–34. Tübingen: Max Niemeyer. [Another version of (1986f). See also (1989e), (1991c).]
- d) "Leonard Bloomfield and the *Cours de linguistique générale*". *Cahiers Ferdinand de Saussure* 43.55–63. [Rev. & expanded version publ. in (1989a: 435–443) and repr. in *Appraisals of Leonard Bloomfield* ed. by John G. Fought (London & New York: Routledge, 1999).]
- e) "The Sources of Modern Sociolinguistics". *The Fifteenth LACUS Forum* ed. by Ruth M. Brend & David G. Lockwood, 343–353. Lake Bluff, Illinois: LACUS. [See also (1986f), (1989c). Rev. & expanded version publ. as (1991c).]
- f) "Towards a History of Americanist Linguistics". *Vingtième Congrès des Algonquistes/Twentieth Annual Algonquian Conference* ed. by William Cowan & Albert D. DeBlois, 1–14. Ottawa: Carleton University Also publ. as (1988c). [Rev. version publ. as (1992d). See also (1994d).]
- g) "Wilhelm von Humboldt and North American Ethnolinguistics: Boas (1894) to Hymes (1961)". *Proceedings of the XX1st Annual Meeting of Cheiron: International Society for History of the Behavioral and Social Sciences, Queen's University, June 15–18, 1989*, 185–192. Kingston, Ontario: Queens University.

1990

- a) (Editor, with G.L. Bursill-Hall and Sten Ebbesen), *De ortu grammaticae: Studies in medieval grammar and linguistic theory in memory of Jan Pinborg*. (SiHoLS, 43.) Amsterdam & Philadelphia: John Benjamins. x + 372 pp.
- b) (Editor, with Hans-Josef Niederehe), *History and Historiography of Linguistics: Proceedings of the Fourth International Conference on the History of the Language Sciences (ICHoLS IV), Trier, 24–27 Aug. 1987*. (SiHoLS, 51.) 2 vols. Amsterdam & Philadelphia: John Benjamins. xx + x + 873 pp.
- c) (Editor, with Henning Andersen), *Historical Linguistics 1987: Papers from the 8th International Conference on Historical Linguistics (Lille, 31 Aug.-4 Sep. 1987)*. (CILT, 66.) Amsterdam & Philadelphia: John Benjamins. x + 577 pp.

- d) (Editor, with Francis P. Dinneen, S.J.), *North American Contributions to the History of Linguistics*. [= *HL* 17:1/2.] Includes a foreword by E. F. K. Koerner. (SiHoLS, 58.) Amsterdam & Philadelphia: John Benjamins. xi + 235 pp.
- e) "Jacob Grimm's Position in the Development of Linguistics as a Science". *The Grimm Brothers and the Germanic Past* ed. by Elmer H. Antonsen, with James W. Marchand & Ladislav Zgusta, 7–23. (SiHoLS, 54) Amsterdam & Philadelphia: John Benjamins. [Rev. version publ. as (1988b). Further expanded version publ. in (1989a).]
- f) "The Place of Friedrich Schlegel in the Development of Historical-Comparative Linguistics". *Leibniz, Humboldt, and the Origins of Comparativism: Proceedings of the international conference held in Rome, 25–28 Sep. 1986* ed. by Tullio De Mauro et al., 239–261. (SiHoLS, 49) Amsterdam & Philadelphia: John Benjamins. Expanded version publ. as (1987f).
- g) "Continuities and Discontinuities in the History of Linguistics". *Proceedings of the 14th International Congress of Linguists (Berlin, 10–14 Aug. 1987)* ed. by Werner Bahner, Jochen Schildt & Dieter Vieweger, III:2649–56. Berlin: Akademie-Verlag. Expanded version publ. in (1989a: 69–78).
- h) "On 'Unrewriting the History of Linguistics'". *History and Historiography of Linguistics: Papers from the Fourth International Conference on the History of the Language Sciences (IChOLS IV), Trier, 24–28 Aug. 1987* ed. by Hans-Josef Nedderhe & E. F. Konrad Koerner, I:63–75. Amsterdam & Philadelphia: John Benjamins. Expanded version publ. in (1989a: 3–12).
- i) "Wilhelm von Humboldt and North American Ethnolinguistics: Boas (1894) to Hymes (1961)". In (1990d), 111–128. Russian transl., "Vil'hel'm fon Gumbol'dt i etnolingvistika v Severnoj Amerike: Ot Boasa do Xajmsa", by Olga I. Romanova, in *Voprosy Jazykoznanija* 41:1.105–113 (Moscow, 1992).
- j) "Ferdinand de Saussure and the Question of the Sources of his Linguistic Theory". *Sprachtheorie und Theorie der Sprachwissenschaft: Geschichte und Perspektiven. Festschrift für Rudolf Engler zum 60. Geburtstag* ed. by Ricarda Liver, Iwar Werlen & Peter Wunderli, 53–66. Tübingen: Gunter Narr, 1990. [Abridged version publ. as (1992g). Rev. version publ. in (1995a: 77–95).]
- k) "Thomas Frank (1925–1990)". *HL* 17:3.421–426. Revised version in *The Virtues of Language: History of language, linguistics and texts* ed. by D. Stein & R. Sornicola, 3–10. (SiHoLS, 87.) Amsterdam & Philadelphia: John Benjamins (1998).

1991

- a) (Editor), Philipp Wegener: *Untersuchungen über die Grundfragen des Sprachlebens*. Including a foreword and analysis of the work by the editor. (CiPL, 5.) Amsterdam & Philadelphia: John Benjamins. xl + ix + 214 pp.
- b) (Editor), *First Person Singular II: Autobiographies by North American scholars in the language sciences*. (SiHoLS, 61.) Amsterdam & Philadelphia: John Benjamins. x + 304 pp.
- c) "Toward a History of Modern Sociolinguistics". *American Speech* 66:1.57–70. [Rev. & expanded version of (1989e); see also (1986f), (1989c). Further rev. version publ. in (1995a: 117–134).]
- d) "History and Historiography of Phonetics: Problems and principles". *Proceedings of LP'90: Linguistics and Phonetics: Prospects and applications (Prague, Aug. 27–31, 1990)* ed. by Bohumil Palek & Přemysl Janota, II:329–344. Prague: Charles University Press. [See also (1978d), (1979b), (1993e).]
- e) "The Sapir-Whorf Hypothesis: A preliminary history". *Newsletter of the Edward Sapir Society of Japan* no.5, 39–55. [Rev. & expanded version publ. as (1992f). See also (1998e & f), (forthcoming e).]
- f) Rev. note on Ulrich Ch.M. Thilo, *Rezeption und Wirkung des Cours de Linguistique générale: Überlegungen zu Geschichte und Historiographie der Sprachwissenschaft* (Tübingen: Gunter Narr, 1989). *Germanistik* 32:1.32–33.
- g) "Saussure, (Mongin-) Ferdinand de". *International Dictionary of Anthropologists* ed. by Christopher Winters et al., 608–610. New York & London: Garland.

1992

- a) (Editor, with Jan Noordegraaf & Kees Versteegh), *The History of Linguistics in the Low Countries*. (SiHoLS, 64.) Amsterdam & Philadelphia: John Benjamins. vi + 400 pp.
- b) (Editor, with Paul Hirschbühler), *The Romance Languages and Modern Linguistic Theory: Papers from the 20th Linguistic Symposium on Romance Languages (LSRL XX), Ottawa, 10–14 Apr. 1990*. (CILT, 91.) Amsterdam & Philadelphia: John Benjamins. viii + 416 pp.
- c) (Associate editor, with Robert H. Robins & Irène Rosier, and with Anders Ahlqvist as main editor), *Diversions of Galway: Papers on the History of Linguistics from ICHoLS VI*. (SiHoLS, 68.) Amsterdam & Philadelphia: John Benjamins. xxvii + 384 pp.
- d) "Toward a History of Americanist Linguistics, with special reference to the

study of Algonquian languages". *Prehistory, History, and Historiography of Language, Speech, and Linguistic Theory* ed. by Bela Brogyanyi, 73–88. Amsterdam & Philadelphia: John Benjamins. [Rev. version of (1988c) and (1989f). See also (1994d).]

e) "William Dwight Whitney and the Influence of Geology on Linguistic Theory in the 19th Century". *Language & Earth: Elective affinities between the emerging sciences of linguistics and geology* ed. by Bernd Naumann, Frans Plank & Gottfried Hofbauer, 271–287. Amsterdam & Philadelphia: John Benjamins.

f) "The Sapir-Whorf Hypothesis: A preliminary history and a bibliographical essay". *Journal of Linguistic Anthropology* 2:2.173–198. [Rev. & expanded version of (1991e). Further rev. & expanded version publ. in (1995a: 203–240). See also (1998e & f), (forthcoming e).]

g) "Ferdinand de Saussure and the Question of the Sources of His Linguistic Theory". *Annali delle Facoltà di Lettere e Filosofia*, Università di Macerata nos. 25/26 (1992–1993), 251–272. Pisa: Giardini. [Abridged version of (1990j).]

h) Rev. note on David Holdcroft, *Saussure: Signs, system, and arbitrariness* (Cambridge & New York: Cambridge University Press, 1991). *Germanistik* 33:1.30.

i) Rev. notes on Gerda Hassler, *Der semantische Wertbegriff in Sprachtheorien vom 18. bis zum 20. Jahrhundert* (Berlin: Akademie-Verlag, 1991) and Renzo Raggiunti, *Philosophische Probleme in der Sprachtheorie Ferdinand de Saussures* (Aachen: Alano/Rader, 1990). *Germanistik* 33:3/4.655, 668–669.

j) "Editor's Foreword". *The Study of Indo-European Vocalism in the 19th Century, from the beginnings to Whitney and Scherer* by Wilbur A. Benware, 2nd printing, p. viii. (SiHoLS, 3.) Amsterdam & Philadelphia: John Benjamins.

k) "Dr. Koerner Responds". *Journal of the History of the Behavioral Sciences* 28:1.67–68. [Reply to Stephen O. Murray's review of (1989a) and (1991b) in the same issue.]

l) Letter to the editors on theory and method of linguistic historiography. *Beiträge zur Geschichte der Sprachwissenschaft* 2:2/3.224.

1993

a) (Editor), Otto Jespersen, *Progress in Language; with special reference to English*. Including a foreword by the editor. Introduction by James D. McCawley (ACiL, 17.) Amsterdam & Philadelphia: John Benjamins. xviii + xii + 370 pp.

b) "The Natural Science Background to the Development of Historical-Comparative Linguistics". *Historical Linguistics 1989: Papers from the 9th International Conference on Historical Linguistics (Rutgers University, Aug. 1989)* ed. by Henk

Aertsen & Robert J. Jeffers, 1–25. (CILT, 106). Amsterdam & Philadelphia: John Benjamins, 1993. [Rev. version publ. as (1995e).]

c) “Chomsky’s Readings of the *Cours de linguistique générale*”. *Proceedings of the XVth International Congress of Linguists, Québec, 9–14 Aug. 1992* ed. by André Crochetière, Jean-Claude Boulanger & Conrad Ouellon, IV:357–360. Sainte-Foy, Québec: Les Presses de l’Université Laval. [Expanded version publ. as (1994c). See also (1993g).]

d) “The Problem of Metalanguage in Linguistic Historiography”. *Studies in Language* 17:1.111–134. Portuguese transl., “O problema da metalinguagem em historiografia da lingüística”, by Maria Cristina F. S. Altman. *D.E.L.T.A.: Revista de Documentação e Estudos em Lingüística Teórica e Aplicada* 12:1.95–124. Rev. version publ. in (1995a: 27–46).

e) “Historiography of Phonetics: The state of the art”. *Journal of the International Phonetic Association* 23:1.1–12 (1993). [Rev. version of (1978d), (1979b), (1991d). Further rev. version publ. in (1995a: 171–202).]

f) Rev. note on Eveline Einhauser (unter Mitarbeit von Theodor Lindken) ed., *Lieberfreund: Die Briefe Hermann Ostoffs an Karl Brugmann, 1875–1904* (Trier: WVT Wissenschaftlicher Verlag Trier, 1992). *Germanistik* 34:4.994.

g) (Organizer and intro.), “Panel Discussion on ‘The History of the Study of the Native Languages of Canada’”. *Proceedings of the XVth International Congress of Linguists, Québec, 9–14 Aug. 1992* ed. by André Crochetière, Jean-Claude Boulanger & Conrad Ouellon, I:215–218. Sainte-Foy, Québec: Les Presses de l’Université Laval. [See also (1993c).]

1994

a) (Editor), Giulio Panconcelli-Calzia: *Geschichtszahlen der Phonetik und Quellenatlas der Phonetik*. With an intro. by the editor. (SiHoLS, 16.) Amsterdam & Philadelphia: John Benjamins. xxxviii + 84 + 88 pp.

b) “Historiography of Linguistics”, “History of Linguistics: Overview”, “Typology and Language Classification: History”. *The Encyclopedia of Language and Linguistics* ed. by R. E. Asher & J. M. Y. Simpson, III: 1570–78, III: 1581–4, IX: 4813–17. Oxford & New York: Pergamon. [Rev. versions publ. in (1995c: 3–6, 1–3, 212–217). See also (1994f)].

c) “Chomsky’s Readings of the *Cours de linguistique générale*”. *Lingua e Stile* 29:2. 267–284. Shorter version publ. in *Lingua et traditio: Geschichte der Sprachwissenschaft und der neueren Philologien. Festschrift für Hans Helmut Christmann zum 65. Geburtstag am 28. Aug. 1994* ed. by Richard Baum, Klaus

Böckle, Franz Josef Hausmann & Franz Lebsanft, 687–699. Tübingen: Gunter Narr, 1994. [Rev. version of (1993c). Further rev. version publ. in (1995a: 96–114).]

d) “*Gramática de la lengua Castellana* de Antonio de Nebrija y el estudio de las lenguas indígenas de las Americas; o, hacia una historia de la lingüística amerindiana”. *Actas del Congreso Internacional de Historiografía Lingüística Nebrija V Centenario, 1492–1992* ed. por Ricardo Escavy, José Miguel Hernández Terrés & Antonio Roldán, vol. II: *Nebrija y las lenguas amerindias*, 17–36. Murcia: Universidad de Murcia. [English version publ. in (1995a: 135–150). See also (1988c), (1989f), (1992d).]

e) “The Anatomy of a Revolution in the Social Sciences: Chomsky in 1962”. *Dhumbadji!: Journal of the Melbourne Association for the History of Language* 1:4.3–17.

f) Biographical entries on Jacob Grimm, M.H. Kruszewski, Hermann Paul, Ferdinand de Saussure, Friedrich von Schlegel, August Schleicher, and William Dwight Whitney. *The Encyclopedia of Language and Linguistics* ed. by R.E. Asher & J.M.Y. Simpson, III: 1505–1506, IV: 1878, VI: 2992–3, VII: 3662–4, VII: 3679–80, VII: 3680–81, IX: 4982–3. Oxford & New York: Pergamon. [See also (1994b).]

1995

a) *Professing Linguistic Historiography*. (SiHoLS, 79.) Amsterdam & Philadelphia: John Benjamins. viii + 274 pp. [Includes (1990j), (1991c), (1992f), (1993d & e), (1994c & d), (1995e & f), (1996b).]

b) (Editor), Mikolaj Kruszewski: *Writings in General Linguistics: On Vocalic Alternations and An Outline of Linguistic Science*. With an intro. by the editor. (ACiL, 11.) Amsterdam & Philadelphia: John Benjamins. xxxix + 188 pp.

c) (Editor, with R.E. Asher), *Concise History of the Language Sciences: From the Sumerians to the Cognitivists*. Oxford & New York: Pergamon. xiv + 497 pp.

d) (Assistant ed., with John E. Joseph & Anders Ahlqvist, and with Kurt R. Jankowsky as chief ed.), *History of Linguistics 1993: Papers from the Sixth International Conference on the History of the Language Sciences (ICHoLS VI), Washington, D.C., 9–14 Aug. 1993*. (SiHoLS, 78). Amsterdam & Philadelphia: John Benjamins. xix + 380 pp.

e) “The Natural Science Model and Linguistic Theory in the 19th and 20th Centuries”. *Săpostavitelno ezikoznanie/Contrastive Linguistics* 19:2.41–59 (Sofia). [Rev. version of (1993b). Further rev. version publ. in (1995a: 7–76).]

- f) "Persistent Issues in Linguistic Historiography". In (1995d), 3–25. [Rev. version publ. in (1995a: 3–26).] Portuguese transl., "Questões que persistem em historiografia da lingüística", by Maria Cristina F. S. Altman, *Revista da Associação Nacional de Pós-Graduação e Pesquisa em Letras e Lingüística* (São Paulo), no. 2, 45–70. Spanish transl., "Problemas persistentes de la historiografía lingüística", by Inés Cuenca Aguilar, *Analecta Malacitana: Revista de la Sección de Filología de la Facultad de Filosofía y Letras, Universidad de Málaga* 19:1.45–66.
- g) Rev. of Randy Allen Harris, *The Linguistics Wars* (New York: Oxford University Press, 1993). *Canadian Journal of Linguistics/Revue Canadienne de Linguistique* 40:2.247–252.

1996

- a) (Editor), Vivian Salmon: *Language & Society in Early Modern England: Essays, 1981–1994*. (SiHoLS, 77.) Amsterdam & Philadelphia: John Benjamins. viii + 276 pp.
- b) "Toward a History of Linguistic Typology". *Linguistic Reconstruction and Typology* ed. by Jacek Fisiak, 1–23. Berlin & New York: Mouton de Gruyter. Rev. version publ. in (1995a: 151–170).
- c) "Schleicher, August". *Lexicon Grammaticorum: Who's Who in the History of World Linguistics* ed. by Harro Stammerjohann et al., 835–836. Tübingen: Max Niemeyer.

1997

- a) "Einar Haugen as a Historian of Linguistics". *American Journal of Germanic Languages and Literatures* 9:2.221–238.
- b) "Linguistics vs Philology: Self-definition of a field or rhetorical stance?". *Language Sciences* 18:2.167–175. [Rev. version of (1982f).] Portuguese transl., "Lingüística e Filologia: O eterno debate", by Cristina Altman & Lineide do Lago Salvador Mosca, *Filologia e Lingüística Portuguesa* (São Paulo) 1:1.4–17.
- c) "On the Linguistic Sources of Roman Jakobson's Work". *Jakobson entre l'Est et l'Ouest, 1915–1939: Un épisode de l'histoire de la culture européenne* ed. by Françoise Gadet & Patrick Sériot, 159–176. Lausanne: Institut de Linguistique et des Sciences du Langage, Université de Lausanne (= *Cahiers de l'Institut de Linguistique et des Sciences du Langage*, no. 9).

d) "Notes on the History of the Concept of Language as a System 'où tout se tient'". *Festschrift in honour of John Hewson* ed. by Vit Bubenik, Marguerite MacKenzie & Harold Paddock, 1–20. St John's, Newfoundland: Memorial University of Newfoundland. [Rev. & expanded version publ. as (1998c).]

1998

a) (Editor), *First Person Singular III: Autobiographies by North American scholars in the language sciences*. (SiHoLS, 88.) Amsterdam & Philadelphia: John Benjamins. x + 267 pp.

b) "History of Linguistics: Attainments and challenges". *Actas del I Congreso de la Sociedad Española de Historiografía Lingüística* ed. by Mauro Fernández Rodríguez, Francisco García Gondar & Nancy Vázquez Veiga, 15–29. Madrid: Arco/Libros. [Transl. into Spanish by Mauro Fernández.]

c) "Noch einmal on the History of the Concept of Language as a 'système où tout se tient'". *Cahiers Ferdinand de Saussure* 51.21–40. [Rev. & expanded version of (1997d).]

d) "Otto Jespersen as a Reader of the *Cours de linguistique générale*". *The Virtues of Language: History in language, linguistics, and texts: Papers in memory of Thomas Frank* ed. by Dieter Stein & Rosanna Sornicola, 151–164. Amsterdam & Philadelphia: John Benjamins. Also publ. as "Jespersen as a Reader of Saussure" in *The Kyushu Review*, 3.1–16 (Fukuoka, Japan, 1998).

e) "On the Immediate and Not So Immediate Sources of the 'Sapir-Whorf Hypothesis'". *Papers from the 1997 Mid-America Linguistics Conference* ed. by Xingzhong Li, Luis López & Tom Stroik, 22–33. Columbia, Missouri: Linguistics Area Program, University of Missouri-Columbia. [See also (1991e), (1992f), (1998f), (forthcoming e).]

f) "Towards a 'Full Pedigree' of the 'Sapir-Whorf Hypothesis': From Locke to Lucy". Paper no. 455 of *Series A: General & Theoretical Papers*, by LAUD Essen, Dept. of English, University of Essen, Germany. Also to appear as (forthcoming e). [See also (1991e), (1992f).]

g) "On the Historiography of the Polish Contribution to the Understanding of Language Change, 1868–1935". *History of Language* 4:2.115–125 (Melbourne).

h) "Wie es eigentlich gewesen — or, Notes concerning the pre-history of the Henry Sweet Society". *Bulletin of the Henry Sweet Society* 30.14–18.

Forthcoming as of 5 Feb. 1999

- a) (Co-editor with Sylvain Auroux, Hans-Josef Niederehe & Kees Versteegh), *History of the Language Sciences: An international handbook on the evolution of the study of language from the beginnings to the present/Histoire des Sciences du Langage: Manuel international d'histoire des études linguistiques des origines à nos jours/Geschichte der Sprachwissenschaften: Ein internationales Handbuch zur Entwicklung der Sprachforschung von den Anfängen bis zur Gegenwart*. 3 vols. Berlin & New York: Walter de Gruyter.
- b) *Historiography of Linguistics: Projects and prospects*. (SiHoLS) Amsterdam & Philadelphia: John Benjamins (1999).
- c) (Editor, with Aleksander Szwedek), *Towards a History of Linguistics in Poland: From early beginnings to the end of the 20th century*. (SiHoLS, 90.) Amsterdam & Philadelphia: John Benjamins.
- d) (Editor, with Hans-Josef Niederehe), *History of Linguistics in Spain/Historia de la lingüística en España, vol. II*. (SiHoLS, 91.) Amsterdam & Philadelphia: John Benjamins.
- e) "Towards a 'Full Pedigree' of the 'Sapir-Whorf Hypothesis': From Locke to Lucy". *Explorations in Linguistic Relativity* ed. by Marjolijn Verspoor & Martin Pütz. Amsterdam & Philadelphia: John Benjamins. [Also publ. as (1998f). See also (1991e), (1992f), (1998e).]
- f) "The Concept of Revolution in Linguistics: Historical, methodological, and philosophical considerations". *History of Linguistics 1996: Selected Papers from the Seventh International Conference on the History of Linguistics (ICHoLS VII)*, Keble College, Oxford, 12–17 Sep. 1996 ed. by David Cram, Andrew Robert Linn & Elke Nowak. Amsterdam & Philadelphia: John Benjamins.
- g) "J. R. Firth and the *Cours de linguistique générale*". *Historical Perspectives on Language Study: Festschrift for Ranko Bugarski* ed. by Olga Mišeska Tomić & Milorad Radovanovich. Amsterdam & Philadelphia: John Benjamins.
- h) "Three Saussures — *un structuraliste avant la lettre*". 'Langue' and 'Parole' in Synchronic and Diachronic Perspective: *Selected proceedings of the 31st annual meeting of the Societas Linguistica Europaea*, St Andrews, Scotland, 26–30 Aug. 1998 ed. by Christopher Beedham. Oxford & New York: Pergamon.
- i) "Ferdinand de Saussure (1857–1913)". *International Encyclopedia of the Social and Behavioral Sciences* ed. by Neil J. Smelser, Paul B. Baltes et al. Oxford & New York: Pergamon.
- j) "Historia de la lingüística: Logros y desafíos". *Analecta Malacitana: Revista de la Sección de Filología de la Facultad de Filosofía y Letras, Universidad de Málaga* 21 (Málaga, Spain, 1999).
- k) "What Is the History of Linguistics Good For?" *Beiträge zur Geschichte der Sprachwissenschaft* 8:3 (1999).

PART IV

Methodological Perspectives

CHAPTER 21

Indo-European Methodology, Bloomfield's Central Algonquian, and Sapir's Distant Genetic Relationships

Regna Darnell
University of Western Ontario

1. Introduction

This paper has emerged from an effort to pin down the precise nature of the considerable influence that Konrad Koerner has had on my longstanding engagement with the intersecting histories of linguistics and anthropology, particularly in North America. Konrad and I have spoken of these matters intermittently since the early 1970s. We collaborated in various contexts during the Edward Sapir centennial celebrations of 1984. More recently, through NAAHoLS (the North American Association for the History of the Language Sciences) and through the various publications he edits so elegantly for John Benjamins, Konrad has challenged me to formulate historiographic problems in Americanist anthropology so that they speak to the concerns of linguists who are not Americanists as well as to those of my anthropological and linguistic colleagues who are.

As a graduate student at the University of Pennsylvania in the late 1960s, I audited historical linguistics with Henry Hoenigswald. Henry welcomed me to his course but apologized in advance that he knew very little about American Indian comparative linguistics and would not often be able to cite examples of direct relevance to my work. That pronouncement seeming to require explanation, he mused for a moment and noted that the genetic relationships involved in American Indian historical work were either so distant that the mere fact of relationship was all that could be ascertained or so shallow in time depth that the details were almost trivially obvious. What interested him, he continued, was the

middle ground (so characteristic of Indo-European) in which the fact of relationship could be taken for granted and detailed history reconstructed (and confirmed by independent evidence). He certainly did not imply that absence of written records might invalidate the comparative method. Nor, however, did he explore the alternative source of evidence, first-hand fieldwork on contemporary languages.

Not surprisingly, most of Henry Hoenigswald's course was about the emergence of Indo-European methodology and results. The promised occasional Amerindian example turned out to be Leonard Bloomfield's Central Algonquian, which I later learned was the sole exemplar for many non-Americanist linguists that Indo-European comparative method could be applied successfully to the study of unwritten languages. These languages are now more appropriately referred to as "exotic" than as "primitive"; both terms, of course, entail inferiority and dramatic alterity.

I am left with some unanswered historiographic questions: What happened to eclipse linguists' memory of the equally rigorous phonetic laws proposed by Sapir for Athabaskan, Uto-Aztekan and Ritwan? Why did Sapir and Bloomfield feel the need to argue that methods derived from the study of Indo-European were applicable to aboriginal languages in North America, a claim which in retrospect seems to have involved flogging a long-dead horse? How was this claim related to the increasing professional autonomy of linguistics in North America in the period following the founding of the Linguistic Society of America and its journal *Language* in 1925? What did all this have to do with the place of linguistics in Boasian anthropology in the early 20th C.?

2. Alternative interpretations of the role of Indo-European in American Indian linguistics

While musing over these matters, I found myself reading for another purpose Lyle Campbell's recent definitive review of the history of American Indian linguistics (1997: 26–89). He explicitly aspires to set to rest what he considers misconceptions in both the historiography and the contemporary practice of American Indian linguistics. His argument foregrounds the continuous and long-standing mutual influences of Indo-European and Americanist linguistics, asserting that "the historical linguistic study of Native American languages was usually up to date with the linguistic methods and theories of the day and not infrequently contributed to them" (1997: 4).

Campbell's extensive documentation is persuasive; indeed, much of it relies on my own work over the past three decades. My 1971 assertion (quoted by

Campbell 1997: 378fn) that Indo-European techniques were not applied to American Indian languages because they were unwritten, that there were no European trained linguists to make such applications, and that Americanist linguistics depended on observation of lexical similarities is certainly an overstatement. I plead only that the context was discussion of the 1891 linguistic classification of John Wesley Powell for the Bureau of American Ethnology, not Amerindian linguistics as a whole. Even so, my reading comes from a quite literal reading of the words of various late 19th and early 20th C. participants about the uniqueness of American developments. John Wesley Powell, for example, was either unaware of or unwilling to acknowledge the roots of his linguistic practices in European comparative method.

At this point, the differences between my standpoint and Campbell's cause us to emphasize different sides of the same coin. I am not suggesting that I am right and Campbell is wrong. Rather, I am engaged by the multiple interpretive possibilities opened up when scholars approach the same historical events and personalities from different angles. The picture is, I believe, more complex than either of us would see working in isolation. Indeed, given that disciplinary history is usually a pretty solitary enterprise, we are fortunate that there is now sufficient literature in the history of Amerindian linguistics to support such an interpretive debate.

1) Campbell is a linguist; I am an anthropologist. He wants to depict a discipline of linguistics located in North America but autonomous from the particular case of American Indian languages, "not merely a Johnny-come-lately stepchild of American anthropology, but rather [with] an independent history of its own" (1997: 28).

I am fascinated by the way this dual disciplinary affiliation plays through the professionalization of American anthropology and linguistics. The dichotomy is perhaps recapitulated in the relatively strong retrospective identification of Leonard Bloomfield with linguistics and of Edward Sapir with anthropology, resulting, at least for some linguists, in the virtual eclipse of much of Sapir's comparativist work.

2) Campbell is interested in relating the search for linguistic family relationships and sound changes to the international history of comparative linguistics. This leads him to focus on pre-professional work by such American scholars as John Pickering, Peter Stephen Du Ponceau, Albert Gallatin, and Thomas Jefferson, who were undeniably in touch with their European counterparts. Even earlier, Roger Williams and Jonathan Edwards identified the Algonquian linguistic family and proposed what we would now recognize as sound correspondences.

I have been interested rather in the professionalization of anthropology (which in North America includes linguistics) and with continuities across the paradigm shift from the Bureau of American Ethnology to Boasian anthropology between 1879 and about 1920 (Darnell 1998). Campbell, in contrast, emphasizes the earlier continuity between Powell and his predecessors, particularly Albert Gallatin.

3) Campbell strongly supports a conservative linguistic classification, correctly noting that Campbell & Mithun (1979) still represents the conservative consensus in American Indian linguistics, confirmed by its use in the languages volume of the *Handbook of North American Indians* (Goddard 1997). Campbell also wants to counter the challenge to this conservative classification by Joseph Greenberg's *Language in the Americas* (1987). In the process, he seems to equate pitfalls inherent in the alternatives of Powell's 58 units in 1891 and Sapir's six in 1921 with those separating Campbell & Mithun's 62 linguistic families in 1979 and Greenberg's three superstocks in 1987. This seems to me to paint with rather too broad strokes of the historiographic brush.

Interestingly enough, the seeds of alternative interpretation are present in Campbell's analysis. It is a question of emphasis.

3. Limitations of the Powell classification

Campbell's interest in the Indo-European linkages of early Americanist work leads him, in my view, to overestimate the linguistic sophistication of the Powell classification. He attributes the drive to achieve a comprehensive linguistic classification of the continent to the Indo-European strain of the tradition Powell inherited, asserting, for example, that it is shared with Thomas Jefferson (1997: 57):

...the methods employed in research on the classification of native languages in the Americas, not surprisingly, were the same as those employed in Europe and elsewhere to establish family relationships and to work out their linguistic history.... American Indian linguistic studies were consistently in tune with developments in European linguistics and Indo-European studies, and frequently contributed significantly to methodological and theoretical linguistic discussions in Europe as well as in America.... The enormous linguistic diversity in the Americas aroused a desire for classification, to bring the vast number of distinct languages into manageable genetic categories (1997: 28).

Campbell (1997: 43–44) quotes Gallatin's 1836 definition of linguistic family as having "at some remote epoch... a common origin... in the same way" as various Indo-European languages. Gallatin, like the linguists on the Bureau staff,

J. Owen Dorsey and Albert Gatschet, was indeed attuned to Indo-European developments. In spite of their Indo-European training, both Dorsey and Gatschet were conservative with regard to genetic relationship, perhaps precisely because the evidence for unwritten languages seemed to them less conclusive in the absence of the written records which could usually be taken for granted in Indo-European. Their work did not, therefore, lead directly to the desired comprehensive classification.

Powell, who needed a classification for ethnological purposes and the applied work on the Bureau mandated by Congress, enthusiastically compared Gallatin's work to that of Linnaeus in biology. Campbell, citing Kroeber (1953), reports (1997: 60) that Powell actually favored biological rather than philological training for anthropology; because his linguists refused to commit themselves on classificatory matters, Powell hired a taxonomic biologist, Henry Henshaw, to transform his extensive vocabulary lists, culled from Smithsonian Institution manuscripts and Bureau fieldwork, into a systematic classification.

There has, in fact, been considerable retrospective discussion of the authorship of the Powell classification. It appeared under Powell's name; the fieldwork was done by various Bureau staff, particularly Dorsey and Gatschet; Henshaw was the primary synthesizer. Campbell (1997: 58) acknowledges that Powell's method was "not very refined", being "a rather impressionistic inspection of rough word lists and vocabularies". He cites William Sturtevant (1959) that the hard part was getting the vocabularies; in Powell's view, anyone could look at them and arrive at the genetic classification. In this light, it is difficult to interpret Powell's methodology — or that applied to the 1891 Bureau classification — as derived from Indo-European.

Campbell's fascinating discussion of Powell's increasing doubts about his ability to distinguish between areal and genetic causes of linguistic similarity (1997: 55, 59) further discredits the philological or comparative character of his reasoning about language classification. Sound correspondences were not part of his linguistic model. Powell's linguistic stocks were intended to establish the fact of relationship within his postulated units. He did not attempt subclassification and apparently thought that the 58 units of the 1891 classification were "equally dissimilar" (Campbell 1997: 62, quoting Darnell 1988). The range from far-flung linguistic families like Algonquian or Athabaskan to linguistic isolates like Zuni, Beothuk, or Kootenay was insignificant in Powell's view of linguistic classification.

Campbell further notes that Powell sided with Otis T. Mason of the United States National Museum against Franz Boas in 1907 about arranging museum exhibits according to evolutionary typology rather than by cultural units (which were at the time usually classified in terms of linguistic families). Boas sought

particular tribal histories, while Mason preferred universals reflecting the broad evolutionary development of "civilization". Here again, Powell's concerns are ethnological rather than linguistic per se. He wants to make sense out of the cultural and linguistic diversity of the continent and seizes upon linguistic classification as the most expedient means to this end.

4. Linguistics as handmaiden to Boasian ethnology

Franz Boas, whose early linguistic work was supported by the Bureau of American Ethnology, came to share Powell's disquiet about his ability to distinguish borrowing from prior genetic relationship among distantly related languages. Unlike Powell, however, Boas acted upon his disquiets, repudiating earlier genetic proposals resulting from his Northwest Coast fieldwork and protesting the more confident reconstructive work of his students, particularly Indo-European-trained Edward Sapir (Boas 1920).

Boasian anthropology implicitly laid claim to an Americanist linguistic tradition indigenous to the continent. Although many of the Boasians were of recent European origin, they challenged an anthropological establishment longer settled on American soil and proud of New World achievements. Campbell (1997: 28) cites "Kroeber's view that Indo-Europeanist methods were too philosophical-typological, too concerned with 'inner form'...whereas the Americanists' methods reflected the practical ethnological expediency of classifying native groups". That is, Kroeber (1913) was not inclined to think that Indo-European methods could solve the problems of North American linguistic classification. Fieldwork and visual inspection were sufficient for practical classification, in Kroeber's case of the linguistic and cultural diversity of the State of California.

Indeed, Kroeber's own classificatory ventures, in collaboration with Roland B. Dixon of Harvard, began with the assumption that they were observing lexical similarities resulting from borrowing, only gradually moving to a genetic interpretation of the same data. Although Campbell (1997: 67) attributes to me the argument that they proceeded "only with reluctance", I think we must also recognize the possibility of dissimulation. This is certainly what Kroeber said. But his motivation included a disinclination to confront Boas's conservatism on classificatory issues; thus, Kroeber emphasized the increasing impossibility of any explanation other than prior genetic relationship for the postulated Hokan and Penutian stocks. The work of Sapir between 1903 and 1913, the period of Kroeber's interpretive shift, buttressed the claims of the comparative method to

elucidate American Indian linguistic data and perhaps enhanced Kroeber's willingness to challenge Boas's conservatism on historical matters.

Despite his acknowledgement of the influence of Hermann Steinthal in his German university years, Boas was a self-taught linguist, learning to describe Eskimo during his Baffin Island fieldwork in the early 1880s. His pedagogy confirmed in his students the assumption that the linguistics necessary for an anthropologist could be learned by trial-and-error. Students took Boas's seminar analyzing his own fieldwork texts from various languages until they understood what was going on (e.g., Lowie 1959). Many of them, who did not aspire to become linguists, produced dissertations consisting of a grammar, dictionary, and texts of various languages which Boas felt in need of immediate salvage. The urgent commitment to record disappearing languages further mitigated against extensive training in comparative Indo-European. (Edward Sapir, the exception among those trained directly by Boas, came to Boasian anthropology *after* his Indo-European training. Pliny Earle Goddard received his Indo-European training from Benjamin Ide Wheeler at the University of California and John Peabody Harrington studied philology in Europe. The rest were amateurs in linguistics as understood independently of anthropology.)

Moreover, Boas was disenchanted with Indo-European because of its evolutionary undertones. Around the turn of the century, Boas's major theoretical project was the critique of evolution. Whereas Campbell (1997: 48) is inclined to deemphasize the evolutionary ethnocentrism of Du Ponceau, William Dwight Whitney, Daniel Brinton, and, indeed, Powell himself (and is undoubtedly correct to do so from a contemporary point of view), I would argue that, for Boas at the time, anything smacking of evolution was apt to be dismissed summarily.

Campbell argues persuasively that the comparativist work during this period of "reductionist frenzy" (1997: 72) centred not around the kind of evidence cited but on the amount of evidence which would warrant a conclusion of proven genetic relationship. He suggests that the debate over distinguishing areal and genetic causes of similarity has resulted in the areal linguistics "now so prominent in most reliable work on proposals of remoter relationship" (1997: 72).

This is, in my view, a peculiarly Americanist methodological tendency, resulting precisely from the application of Indo-European method in the absence of independent confirmation from written records. This methodology was formulated elegantly by Sapir in *Time Perspective in Aboriginal American Culture: A study in method* (1916). This is perhaps the most extreme example in Boasian anthropology of linguistics as handmaiden to ethnology, and it came from the Boasian who was most directly an Indo-Europeanist. Sapir argued to his colleagues in ethnology that linguistic method could distinguish genetic from

areal relationship because sound changes were by definition characteristic only of the former. The rest of culture was patterned in a way which failed to preserve historical development in contemporary forms distinguishable as to common origin or borrowing.

Campbell does not emphasize Sapir's reliance on Indo-European methods in this context, presumably because it led Sapir to what he considers unjustified genetic hypotheses. Campbell cites Kroeber (1940) that Sapir's six-unit classification of 1921 was not scientific in its method (quoted 1997: 75):

From one point of view such a procedure is nothing less than forecasting... It is in no sense whatever a definable or controllable method of science or scholarship. The danger... is that its prophecies may be mistaken especially by non-linguists for proved or probable findings. Tremendous havoc can be worked when archaeologists or ethnologists begin to build [sic in Campbell] structures of inference on Sapir's brilliant but flimsy gossamer web of prophecies as if it were a solid foundation.

There is no discussion of Kroeber's turn away from linguistic classification in the intervening years of his career. Indeed, Campbell, still relying on Kroeber's repudiation decades later of his own complicity with Sapir's drastic reduction of the number of linguistic stocks in North America, concludes (1997: 76): "Later, it came to be assumed that this classification [Sapir's six-units] had been established by legitimate linguistic methods, and thus it became entrenched in the literature".

After further detractions from the conclusions of the six-unit classification, Campbell (1997: 76) implies that Sapir, later in his life, implicitly recanted by citing [in an article not devoted to his applications of the comparative method] large numbers of unrelated linguistic stocks in North America. It is certainly the case that Sapir turned away from dramatic reduction of the number of linguistic families in the continent after presenting his own classification. He had, after all, had his say. It is also the case that Americanist linguistics was less obsessed with classification after the 1920s. Where I am uncomfortable, however, is with the conclusion that Sapir changed his mind about the ability of genetic relationship to persist identifiably at great time depth and about the utility of postulating distant relationships even without complete or fully convincing evidence. Even in the six-unit classification, Sapir allowed for an intermediate level acceptable even to the more conservative among his contemporaries. There was room for further research to clarify the proposed more distant relationships.

In any case, when Campbell turns to Bloomfield (1925, 1928, 1946) for demonstration that sound change is regular in unwritten or exotic languages (1997: 77), it is difficult to escape the conclusion that Sapir's application of

Indo-European method to both demonstrable and only suggestive cases of putative linguistic relationship is taken to have tainted all of his comparativist work. This conclusion seems to me perverse. I am not prepared to dismiss the legitimacy of Sapir's linguistic methods in a cavalier fashion.

5. How Bloomfield came to provide the exemplar

In 1931, Sapir was invited to write about methods of linguistics for a case book in the social sciences, representing the successes of the interdisciplinary synthesis that crystallized around the Chicago School of Sociology in the 1920s. With a rare and undoubtedly motivated display of public modesty, Sapir made his point about the special properties of the comparative method for studying American Indian languages with reference to the work of Leonard Bloomfield.

As his point of method for the volume, Sapir stressed (1949 [1931]: 73) that Bloomfield's contribution was to have "tested" the Indo-European-derived "concept of phonetic law" in relation to what were then called "primitive" languages. Such phonetic laws were, in Sapir's view, foundational to "the scientific study of language".

Sapir's initial illustrations of regular sound change were conventional ones taken from Indo-European. But he went on to assert that evidence was rapidly accumulating, presumably through Boasian fieldwork, that the same methods could be applied to any languages: "If these laws are more difficult to discover in primitive languages, this is not due to any special characteristic which these languages possess but merely to the inadequate technique of some who have tried to study them" (1949 [1931]: 74). He was explicit regarding the applicability of European linguistic models in America: the methodology of Bloomfield's Proto-Central Algonquian reconstruction "is precisely the same as the methodology which is used in Indo-European linguistics" (1949 [1931]: 75). Sapir also implied that many of his colleagues, with or without training in Indo-European, lacked the historical imagination to trust sound changes as reliable evidence for distant relationships among American linguistic families (cf. Darnell 1990). Yet his reliance on Bloomfield to make the case about the regularity of sound change, in general and in American language families in particular, suggests that he may have wanted to associate himself with narrowly comparative reconstruction rather than with his own more speculative proposals over the previous two decades.

Sapir illustrates five phonetic laws from Bloomfield (1925), particularly praising a sixth law on the basis of a single set of correspondences, noting that later information from his own fieldwork on Swampy Cree confirmed the analysis,

as well as the predictive capacity of the comparative method (Bloomfield 1928).

Only after considerable discussion of Bloomfield's work on Central Algonquian did he turn to his own work (1949 [1931]: 78): "Bloomfield's experience with the Central Algonquian dialects is entirely parallel to my own with the Athabaskan languages".

Nonetheless, Sapir chose his historical problems in a different way than did Bloomfield. Central Algonquian was a tidy geographically contiguous unit, wherein genetic relationship had long been recognized. Sapir, in contrast, was fascinated with Athabaskan partly because of its irregular distribution in North America. Ritwan was intriguing because its distant relationship to Algonquian seemed geographically anomalous. Uto-Aztekan was puzzling because of the range of cultural development separating Nahuatl, the language of the Aztec civilization, from the "root-digging savages" of the southwestern desert.

Sapir's Athabaskan sound patterns were presented in terms of three "representative" languages: Navajo, Chippewyan, and Hupa. He emphasized "pattern" (that is, phonemic contrast) rather than particular sounds. His form of presentation is fully parallel to that used for Bloomfield's Central Algonquian. We may infer that Sapir wanted to claim for his own comparative work the rigor of the Indo-European method as practised by Bloomfield. He must have realized that his six-unit classification was misunderstood (in his view) by many of his more narrowly focused linguistic colleagues. In this paper, he went on the offensive.

Sapir was not the only Americanist linguist to acknowledge the exemplary status of the Central Algonquian reconstruction. Charles Hockett (1964 [1948]: 599) saw "the descriptive and comparative structure of Algonquian" as "one of Leonard Bloomfield's life works". Nonetheless, those linguists who praised it usually did so on "indirect evidence" derived from Bloomfield's stature as a Germanicist: "since Bloomfield's other work proves him a sound scholar, his Algonquian studies must be sound too. Algonquian, after all, is an out-of-the-way language family, and few have concerned themselves with it". Hockett presents sixteen principles of the comparative method illustrated in Bloomfield's Algonquian work, optimistically concluding (1964 [1948]: 609):

Yet if, for any one of the sixteen, Bloomfield's Algonquian evidence stood alone — if there were no comparative Germanic, comparative Romance, comparative Indo-European, comparative Semitic, and so on — that isolated support for the principle would still be persuasively solid.

Hockett does not directly address the implicit ethnocentrism of European linguistics, particularly its assumption that the important languages of the world are those with written traditions. Regardless of the quality of Bloomfield's

reconstructions, for many non-Americanist linguists they could not, in principle, hold the paradigmatic potential of the classic Indo-European cases from which the comparative method was derived. During the interwar years, even in North America, Americanist linguists still had to defend the importance of work on unwritten languages and the validity of their applications of Indo-European methods to them. Otherwise, Sapir, Bloomfield, and Hockett would have spent less time making the case.

This was particularly crucial with the professionalization of linguistics in North America, since Americanist linguists found themselves engaged with colleagues whose priorities were quite different and whose attention to American Indian languages was perfunctory at best. It is unclear how many of Sapir's colleagues agreed with him in this period that "When it comes to linguistic form, Plato walks with the Macedonian swineherd, Confucius with the head-hunting savage of Assam" (1921: 234).

Hockett himself stressed, in retrospect, that the methods of anthropological linguists [i.e., fieldwork] were unfamiliar to the philologists and comparativists. The Americanists could not draw on "inherited bodies of texts" as Bloomfield had been trained to do in Germanics. Rather, "he transferred the [Indo-European] method bodily — as also, indeed, did Boas, Sapir and every other investigator of such languages ... since there were no texts, they had to be gathered by writing down accounts from the lips of native speakers" (1989: 5). Then and only then could the anthropological linguist proceed philologically, using the comparative method.

Hockett (1989: 4) continues to assert that Bloomfield's Central Algonquian, with "reconstructions ... as full, realistic and convincing as any in Indo-European", demonstrates the absolute universality of Neo-Grammarian regular sound change. Hockett argues that a single counterexample would render the comparative method a "mistake". This sense of testing the comparative method against data from widespread and historically independent linguistic families does not seem to me to ever have characterized linguistics overall, in Europe or North America. Although the equivalence in principle of all human languages can no longer be denied, it has been particularly clear over the last four decades that not all linguists turn in practice to the study of the widest possible sample of such languages.

There remains a fundamental discontinuity between American linguistics and European. Hymes & Fought (1975: 917) cite "the inhospitability to European theory" of various American structuralist linguists, in part because European refugees were getting jobs during the Depression while American linguists remained unemployed. Post-war linguistics was even more isolationist and self-consciously American. Hymes & Fought conclude that European models are not

needed to account for the grammatical and comparative work of Sapir and Boas. What they find particularly fascinating is that (1975: 952):

...in the United States... the scholars who participated in the founding of structural description participated also in the founding of the comparative-historical linguistics of a great many language families. Comparative philology was effectively introduced to these families by structuralists — Bloomfield in Algonquian, Sapir in Uto-Aztecan, Athabaskan, etc.

Indeed, the whole Americanist emphasis on grammatical categories is not present in the European tradition of the same period. The synthesis of synchronic and diachronic analysis arose in good part from the local exigencies of Amerindian fieldwork. In this Americanist tradition, Indo-European was only one of the methods employed.

6. Conclusion

In spite of a long history of comparative linguistics of American Indian language families, both Leonard Bloomfield and Edward Sapir found it necessary to argue in the 1920s and 1930s that Indo-European methods could be reliably applied to the study of unwritten languages; furthermore, this application was taken to demonstrate the universality of the method itself. In the view of its most prestigious practitioners, American Indian linguistics had much to contribute to general linguistics. Both Edward Sapir and Leonard Bloomfield were adamant that the autonomous American discipline of linguistics would apply itself to the study of all languages.

Hymes & Fought define a “first Yale school” of linguistics which emerged around Sapir in the 1930s, a decade before the widely acknowledged Yale school associated with Bloomfield thereafter. The principles of Sapir’s linguistics (1975: 997) begin with the application of structuralist methods “to both exotic and well known languages”. The professionalization of linguistics and the rescue of what we now call endangered languages are the next foundation points. These three principles do not distinguish Sapir from Bloomfield. Sapir and Bloomfield shared a concern with genetic relationship of linguistic families, although they had somewhat different notions of how this should be approached. Both also wanted to tie linguistics to other disciplines and to practical matters, for example, education; again, however, their approaches were quite different in practice.

The substantial overlap in goals suggests, however, that Sapir and Bloomfield deployed the methods of Indo-European to similar ends and with similar rigor. Sapir’s Athabaskan, Uto-Aztecan and perhaps Ritwan have stood the test

of time, alongside Bloomfield's Algonquian, despite the failure of Sapir's work to attain the same paradigmatic status for general linguistics. The major difference between the two efforts, then, appears to be that Sapir also attempted more speculative reconstructions which have been embraced by ethnologists and archaeologists, rejected or reformulated by linguists, especially those outside anthropology.

On Ritwan, for example, Goddard (1986) argues that Sapir was correct in linking California Wiyot and Yurok to Algonquian, although sometimes for the wrong reasons. Goddard would have been happier with fewer correspondences, each unassailable. Sapir, however, preferred to list all possible evidence, noting that some cognates were uncertain. He argued that, in the long run, this would allow future researchers to correct and expand on his work.

The continuity of Indo-European influences on the study of American Indian languages as accessible to contemporary historiography, however, must be distinguished from the perceptions of actors at any given moment of that history. Certainly neither Sapir nor Bloomfield evidenced extensive knowledge of prior work (extensively discussed by Campbell 1997) which in retrospect seems foundational. The Boasian/Americanist emphasis on fieldwork may well have foreshortened this history of the application of Indo-European methodology to the study of American Indian languages.

Author's address

Regna Darnell
Department of Anthropology
University of Western Ontario
London, Ontario, CANADA N6A 5C2
rdarnell@julian.uwo.ca

References

- Bloomfield, Leonard. 1925. "On the Sound-System of Central Algonquian". *Language* 1. 130-156.
———. 1928. "A Note on Sound Change". *Language* 5.99-100.
———. 1946. "Algonquian". *Linguistic Structures of Native America* ed. by Harry Hoijer, 85-129. New York: Viking Fund.
Boas, Franz. 1920. "The Classification of American Languages". *American Anthropologist* 22.367-376.

- Campbell, Lyle. 1997. *American Indian Languages: The historical linguistics of Native America*. Oxford: Oxford University Press.
- & Marianne Mithun, eds. 1979. *The Languages of Native America: Historical and comparative assessment*. Austin: University of Texas Press.
- Darnell, Regna. 1990. *Edward Sapir: Linguist, anthropologist, humanist*. Berkeley & Los Angeles: University of California Press.
- . 1998. *And Along Came Boas: Continuity and revolution in Americanist anthropology*. Amsterdam & Philadelphia: John Benjamins.
- Goddard, Ives. 1986. "Sapir's Comparative Method". *New Perspectives in Language, Culture and Personality* ed. by William Cowan, Michael Foster & Konrad Koerner, 191–214. Amsterdam & Philadelphia: John Benjamins.
- , ed 1997. *Handbook of North American Indians, volume 17: Languages*. Washington, D.C.: Smithsonian Institution.
- Greenberg, Joseph. 1987. *Language in the Americas*. Stanford: Stanford University Press.
- Hockett, Charles F. 1964 [1948]. "Implications of Bloomfield's Algonquian Studies". *Language in Culture and Society* ed. by Dell Hymes, 599–609. New York: Harper & Row.
- . 1989. "Leonard Bloomfield: After fifty years". *Yale Graduate Journal of Anthropology* 2, 1–11.
- Hymes, Dell & John Fought. 1975. *American Structuralism. Current Trends in Linguistics 13: Historiography of Linguistics*. The Hague: Mouton.
- Kroeber, Alfred. 1913. "The Determination of Linguistic Relationship". *Anthropos* 8, 389–401.
- . 1940. "Conclusions: The present status of Americanist problems". *The Maya and Their Neighbors* ed. by Clarence L. Hay et al., 460–487. New York: D. Appleton-Century.
- . 1953. "Concluding Review". *An Appraisal of Anthropology Today* ed. by Sol Tax, 357–376. Chicago: University of Chicago Press.
- Lowie, Robert H. 1959. *Robert H. Lowie, Ethnologist*. Berkeley & Los Angeles: University of California Press.
- Powell, John Wesley. 1891. *Indian Linguistic Families of America North of Mexico. Seventh Annual Report of the Bureau of American Ethnology*: 1–142. Washington, D.C.: Government Printing Office.
- Sapir, Edward. 1916. *Time Perspective in Aboriginal American Culture: A study in method*. Ottawa: Canada, Department of Mines, Geological Survey, Memoir 90, Anthropological Series 13.
- . 1921a. "A Bird's Eye View of American Languages North of Mexico". *Science* 54.408.
- . 1921b. *Language: An introduction to the study of speech*. New York: Harcourt, Brace & World.
- Sturtevant, William C. 1959. "The Authorship of the Powell Linguistic Classification". *International Journal of American Linguistics* 25.196–199.

CHAPTER 22

The Need for Phonetically Accurate Notation in the General Progress of Linguistics

Saul Levin

State University of New York at Binghamton

The friendship and gratitude that I feel toward Konrad Koerner are here expressed in an essay dealing with a field where my research approaches his. While I cannot do justice, as he does, to the history of how linguistic science has developed, my particular studies of languages have at times led me to look into the achievements of certain scholars in the past, which conduced to an amelioration of grammatical knowledge, or at least were so intended. The difficulties that they grappled with were fruitful, for from them emerged a somewhat improved terminology or notation, or both.

1. Greek and Latin as school languages

My starting point will seem mundane, or even trivial — the attachment, early in the Christian era, of an explanatory adjective to the names of four Greek letters: Ε ψιλόν “bare” or “plain”, Υ ψιλόν, Ο μικρόν “little”, and Ω μέγα “big”.¹ In the process of pedagogy the adjective merged to form a trisyllabic name² and ousted the original name, which had consisted of just the vowel sound itself with only the phonological adjustment to make it a normal, pronounceable word: [e:, hy:, o:, ɔ:]. As the former long vowels [e:] and [o:] had shifted in quality to [i] and [u] respectively (and were written as digraphs, EI and OY), the old names in their modified pronunciation were misleading to beginners, making it harder for them to learn the function of the simple vowels E and O. The term ψιλόν, however, addressed something else: the tendency for AI δίφογγον “with two sounds” to be pronounced the same as E, simply a front-vowel of intermediate aperture — and similarly for OI δίφογγον to be pronounced the same as Y (like *ü* in German).

For the purpose of holding on to CLASSICAL Greek, the grammarians and the teachers they were presumably writing for had to make a sharp distinction, at least in orthography if not in pronunciation too. Otherwise their pupils were liable to confuse the ending of the second person plural active, as in ΑΥΕΤΕ "you are untying" (which is also imperative) with that of the third person singular middle ΑΥΕΤΑΙ "he/she is getting untied" (or "getting someone untied"), and to make many other mistakes.³ The effort succeeded to this extent: the WRITTEN IMAGE of classical Greek was preserved with admirable fidelity,⁴ although eventually no scholar remained in the Byzantine world who still kept the diphthongal pronunciation.⁵

In the long run, the very zeal of the Byzantines for orthography enabled some Hellenists in western Europe to recapture more and more of the real but lost sounds of ancient Greek. From Erasmus on (around 1500), the scholars of other nations who learned classical Greek realized that the current pronunciation of the Greek teachers themselves had leveled several ancient distinctions, mainly between vowels. For by that time the phonology of Greek was down to five simple vowels /i e a o u/, with no distinction of length and no diphthongs; and so the learned among the Greeks simply applied this vernacular phonology when they read Homer, Sophocles, Plato, Demosthenes, and other authors of roughly two thousand years earlier, and when they conversed in this classical school language. Erasmus' most convincing argument against the modern Greek pronunciation of the ancient classics was drawn from the Latin renderings in the age of Cicero (1st century B.C.). A transcription such as *Phoenixe* for φοινίκη proves that the pronunciation of both οι and η as [i] must have developed subsequently. Therefore the Erasmians (as the resentful and defensive Greeks called them) restored η to equivalence with the Latin long *e*, as well as restoring the diphthongal pronunciation of οι and αι.

It took much longer to reform the school pronunciation of the Latin diphthongs *oe* and *ae*, which had been monophthongized in the early centuries of the Christian era and eventually simplified to *e* in the spelling of the medieval scribes. Although scholars who traveled were bound to notice the diverse national pronunciations of Latin, those diphthongs were not retained anywhere. The humanists of Italy, and of other countries later, succeeded best in reclaiming the VISIBLE heritage of imperial Rome, which included the spelling of Latin words;⁶ but the ancient sounds were more elusive.

The great Renaissance produced inscriptions in capital letters on new or renovated churches and palaces and on tombstones; but it was impractical to do without the medieval minuscule letters in books.⁷ The eventual result was a regulated mixture of capitals and minuscules with a further complication: The

reduced capital *v*, while serving for many generations as a purely graphic alternative to the rounded minuscule *u*, was preferred by the scribes only at the beginning of a word, where this twentieth letter of the alphabet — called [u:] both in ancient times and in most of the national pronunciations of Latin⁸ — functions more often as a consonant. That gave Petrus Ramus (Pierre de la Ramée, 1515–72) the idea of redefining the *v* as strictly a consonant, a voiced labial fricative; he did not realize that it deviated from the ancient labial semi-vowel, nor did his contemporaries. So his reform, calling for *venire* “to come” and *revenir* “to come back” (where the scribes used to write *reuenire*) but *una* “one” (fem., where they had favored *vna*), unfortunately obfuscated the authentic phonology of Latin; but it brought out well the difference in French between *venir*, *revenir* on the one hand and *une* on the other. It also improved the orthography of nearly all the rest of the languages of Europe that are written in the Latin alphabet.

A faulty pronunciation did not always preclude mastery of the classical language otherwise, so as to compose a new text with correct morphology, syntax, and idiom. Some scholarly poets — Petrarch and Milton among them — even turned out quite lengthy Latin poems, not far below the quality of their finest poetry in Tuscan or in English.⁹ The confusions that distorted the phonology of Latin, as taught in the schools, were finally sorted out by Corssen (1858–59) and his successors at the universities of Germany. To put their findings into practice in the classroom, it took their American graduate students, coming home with the prestige of a doctorate from a German university, and appreciating the greater need to reform the current school pronunciation of Latin than was felt in Germany. For, to cite an extreme example, *sīcae* “daggers” (nominative) had come to be pronounced [s’aisi:] in England and America, which reversed the vowels of the two syllables and altogether neutralized the difference between the consonants; in Germany it was [zī:tse:], maintaining both a consonantal and a vocalic distinction, though at a considerable remove from true ancient Latin.

The reform made rapid headway in the United States and was afterwards embraced by open-minded teachers in Great Britain.¹⁰ Since then it has gained more slowly on the Continent. It is most strongly opposed in Italy; for it would cut across certain orthographic conventions of Italian: e.g., it requires pronouncing the Latin *dīcō* “I say” and *dīcit* “he/she says” with the same velar plosive [k], whereas the vernacular *dice* “says” has instead the affricate [č] while *dico* maintains the original [k]. To do justice to both the ancient and the modern language, one must get past the ambiguities of spelling.

So far I have treated the concerns of scholars coping with the weaknesses in both the oral and the written tradition of a classical language that was losing,

or had quite lost, its former vernacular base. While still propagated through schooling, it was bound to deteriorate from the classic norm, unless an extraordinary effort were made to correct the deviant tendencies.

2. Two separate advances, Indo-European linguistics and the International Phonetic Alphabet

More broadly, the linguistic profession, during the nineteenth and twentieth centuries, has in principle accepted the importance of describing the sounds of each language precisely. A major upsurge began with the addition of Sanskrit to the framework of linguistic data already familiar to the scholars of Europe. Thanks to the superior acumen of the Brahmins, the structure of this language emerged with amazing analytic clarity from the texts which they had transmitted from an age even earlier than the *Iliad*. But the Nagari script, being syllabic rather than alphabetic, was troublesome to European scholarship.

So, contrary to the precedents of citing Greek or Hebrew or Arabic words in the original script, the Occidental grammarians would transcribe Sanskrit words into the roughly equivalent Latin letters. While this facilitated the recognition of cognates such as

ājanti (अजंति) “they drive” : Greek ἄγουσι (Doric -οντι) and
ājata (अजत, imperative pl.) : Greek ἄγετε

it also encouraged a mistaken identification of the Sanskrit vowel, transcribed *a*, with the Greek *a*. To be sure, it is unlike the Greek front-vowel *ε* and the back-vowel *ο*; but neither is it wide-open like *α*. Instead it is like the English vowel in *punch* and occurs far more often in Sanskrit words than any other one sound. But this perfectly intermediate vowel is absent from Greek and Latin, as well as many modern European languages, including German; and German scholars were the ones that led the way in comparative linguistics, studying Sanskrit a lot more than the British did. They managed to get this sound right when learning English; but since they approached Sanskrit as a book language, they did not bother to differentiate the same Sanskrit vowel from their German [a]. And obviously they could not write it *u*, as the English did in words from Hindi such as *pundit* (< Skt. पंडितः).

A satisfactory grapheme [ʌ] for this noteworthy vowel was devised as part of the International Phonetic Alphabet around 1900. The inventors, Paul Passy and his colleagues, were Frenchmen, seeking an effective graphic approach to

instruct their countrymen in the pronunciation of English.¹¹ By that time, as more and more Frenchmen perceived the value of learning English, the utter discord between spelling and pronunciation in both languages at last motivated a well thought-out method for surmounting it. Over the many previous centuries, when French was widely taught in England, the pupils had gone virtually unaided by any written clues to the discrepant pronunciation of the vowel letters in *retient* "(he/she) keeps" / *reliant* "(they) tie down", *faisons* "let's make" / *faisans* "pheasants", *soignons* "let's take care" / *oignons* "onions", *eu* "had" (participle) / *jeu* "game", and of the consonant letters in *vill*e "city" / *fil*le "daughter", *dot* "dowry" / *so*t "fool", *os* "bone" / *dos* "back", etc.

Pedagogical experience has shown that the IPA, with its numerous unfamiliar characters, is too subtle for many persons of ordinary literacy; but for linguists it is a boon indeed, adaptable more or less readily to every language in the world. The Sanskritists and the Indo-Europeanists have yet to catch up; they still do not write [ʌ̌janti] for अ जं ति or [ʌ̌jAtA] for अ जं त.

To illustrate not only the importance of exact phonetic data but also how elusive they can be, I would mention the endless controversy, within Indo-European linguistics, between the proponents of prehistoric laryngeal consonants and their opponents. Insofar as those consonants are posited in order to account for odd correspondences between the Greek and the Sanskrit vowels, as in the root syllable of the verbal adjectives

θετόν:	धि त म्	[dʰitám] ¹²	"set, put"
στατόν:	स्थि त म्	[stʰitám]	"standing"
δοτόν:	-त्त म्	[-ttám]	"given",

we can scarcely expect an exact phonetic description of the hypothetical consonants; rather, the mere cover-symbols, *H₁, *H₂, *H₃ would seem most prudent.¹³ But we have definite information, from two quite separate Indo-European sources, of a consonant at the beginning of words whose cognates in other Indo-European languages begin with a vowel:

(1) The decipherment of Hittite, that most ancient language, written in the Sumero-Akkadian syllabary, revealed many words with an initial {ḫa-}, {ḫe-}, {ḫi-}, or {ḫu-}; {ḫa-an-ti} "in front, separate", cognate to Sanskrit अं ति [anti] "in front, near", Latin *ante* "in front", Greek ἄντι; "instead of", etc. A phonetic definition of this Hittite consonant can only be indirect and tentative, since Akkadian and Sumerian were also long extinct, forgotten languages, recovered

through decipherment. The Semitic cognates of Akkadian words with {h} have a velar fricative; e.g., Arabic خَنَقَ {ḫanaqa} "he choked" (likewise in Geʿez): Akkadian {ḫa-ni-iq} "choking". This is a velar fricative, like the Cyrillic letter x in Church Slavonic ходити, Russian ходить "to go" (cf. German *auch*).

(2) Several Germanic languages have a pronounced but unwritten glottal stop [ʔ] before the written initial vowel. This is quite audible in standard High German: [ʔ] *Antwort* "answer", [ʔ] *Ende* "end", etc. That it was more general in early Germanic is inferred from the rules of 'alliterative' versification in Old English and Old Saxon: Words written with any initial vowel are treated as if it were preceded by one and the same unwritten consonant; e.g.,

sawle secan þone syscaðan

ænig ofer eorþan irenna cyst

"to seek the soul, that sin-scathed one, any choice one of the irons on earth [would not touch]" (*Beowulf* 801-802).¹⁴

Not all Germanists have agreed that *eorþan* and the like must have been pronounced with [ʔ-]; but those who deny it are obliged to regard the modern High German [ʔ-] as a separate recent development, seemingly out of nowhere.

This becomes all the more incredible when we bring in the Semitic cognates of certain Germanic (and Indo-European) words, such as

Arabic اَرْضَ {ʔarḍun} (nom.; gen. اَرْضِي {ʔarḍin}, accus. اَرْضًا {ʔarḍan})
German [ʔ] *Erde(n)*¹⁵

It would make no sense to isolate the actual German [ʔ] from the same consonant in Semitic on the theoretical ground that prehistoric Indo-European lacked this sound. The words in question are distributed unevenly among the Indo-European languages; so it may be unnecessary to posit a proto-Indo-European [ʔ], subsequently lost in the precursors of Greek and most other branches of the family.

It remains obscure what prehistoric connection, if any, there may have been between the quality of the vowel in θε-, στα-, δο- and a particular laryngeal (or guttural) consonant. When the Ugaritic tablets, written in alphabetic cuneiform characters, were deciphered in the 1930s, three letters were found to correspond to the glottal stop ʔ of Hebrew and Aramaic (ʔ in Arabic), and the three are correlated with a central, a front, and a back vowel of the well preserved Semitic languages; hence the transliteration {ʔ^a, ʔⁱ, ʔ^u}, the weak glottal consonant being colored (so to speak) by the adjacent vowel.¹⁶

An Ugaritic abecedarium that was dug up begins with {ʔ^a}, in the same position as ʔ in the Phoenician-Hebrew alphabet.¹⁷ {ʔⁱ} and {ʔ^u} are among the

added letters at the end of the Ugaritic abecedarium, which stand for sounds not present or not recognized in Phoenician. The graphic divergence, within these neighboring and closely related Semitic languages, may be best explained thus: the very meagerness of phonetic features in the consonant [ʔ] makes the adjacent vowel more noticeable, if not in the absolutely consonantal Phoenician alphabet, then at least in the Ugaritic, which was somewhat open to influence from the Akkadian cuneiform syllabary.

3. Phonemic as distinct from phonetic notation

On the whole, the data in written form which a linguist has to deal with are adequate, though falling short of ideal accuracy. The experience of research has often led to improvements in notation, among them the oblique brackets / / to enclose the phonemes of a word, as in /tæk, stæk/ to represent the English words *tack*, *stack*; the older device of phonetic brackets [] ought to record the sub-phonemic aspiration in [t^hæk, stæk^h]. Further analysis, however, brings out finer and finer phonetic details, such as the relatively long vowel before the voiced consonant in *tag*, *stag*, and regional differences in the quality of the vowel; e.g., it becomes a diphthong [æ^e] or [æⁱ] in the southeastern United States.

Within linguistic science the purposes of phonetic notation are multiple. But in each case it should enable the reader to pronounce the word or phrase just as the author intended, and thereby to grasp the author's linguistic argument. Accordingly the notation ought to be free from the ambiguities, as well as redundancies, that complicate the orthography of nearly every European language. The International Phonetic Alphabet, while first of all targeting English and French because of their grossly unphonetic spelling, was meant to be applied consistently to all languages; e.g., [y] in brackets should stand only for the close rounded front vowel (written *u* in French and normally *ü* in German), never for a consonant.

The rules of the IPA, however, amount to little more than recommendations; they cannot be enforced like the standard spelling of a language taught in elementary school. Individual linguists use the IPA at their own discretion and with numerous modifications; for they are accustomed to many languages that not only lack the sound designated by a certain IPA character but employ the very same letter — with internal consistency — for a quite different sound. In discussing Spanish there is little need for the IPA, and some risk of confusion if you bring [j] for the semi-vowel into words like *peyor* “worse” and *mayor* “greater”, since the letter *j* serves as in *mejor* “better”, *joven* “young” for the

voiceless velar fricative. This consonant, according to the IPA, is [x] — a rather arbitrary decision of the planners. They chose it on the basis of the FORMER Spanish spelling of certain words; e.g., *dixe* “I said”, *exe* “axle”, where the Latin *x* [ks] had changed first to [ss] and then underwent palatalization. But after both this [j] and the voiced affricate [dʒ], written *j*, merged into the voiceless velar fricative, the Real Academia Española reformed the spelling to use only *j* for the recently developed sound, and restricted the letter *x* to learned borrowings, mainly from Latin (e.g., *exacto*, *complexo*), which maintain the ancient consonant group [ks].¹⁸

The masterminds of the IPA were unwise to assign an unfamiliar value to some familiar letters. They did not want to use diacritical marks more than necessary; but subsequent experience has shown that [i̯] serves better than [j], and [h̥] better than [x]. For the voiceless and the voiced palatal sibilants they ought to have taken over *š* and *ž* from Czech orthography, rather than drawing upon *f* and *z* (which were minuscule variants of the letters S and Z in an obsolete style of medieval penmanship).

The IPA has had little effect upon the transliterations previously in use among scholars for languages not written in the Latin alphabet. As I pointed out above, the Sanskrit vowel customarily transcribed *a* is really [ʌ]; a structural analysis of the language would be definitely clarified by replacing the inaccurate *a*, and so would the relation of Sanskrit to cognate languages.

On the other hand, the three Arabic vowels that are customarily transcribed *a*, *i*, and *u* are perfectly appropriate on the phonemic level, but each of them has a wide phonetic range of actualization, depending upon the consonantal environment. A phonetically useful transcription would require quite a few more of the IPA vowel letters; the Arabists have not yet found that worthwhile.

The study of Biblical Hebrew suffers from the reluctance of most Hebraists to put aside an inveterate misunderstanding of the vowel notation. It was devised admirably by the expert readers in the city of Tiberias about 1200 years ago; however, the subsequent grammarians, Jewish and afterwards Christian, confused a system that was intrinsically clear. Their fallacies have been refuted, but not enough to break through the wall of incorrect pedagogy. Nearly all students of this ancient language are still taught to mispronounce it; few of them ever learn to surmount the lessons of their misguided teachers. The vowels of the IPA, including [ɛ] and [ɔ], suit Hebrew quite neatly; but the failure of Hebraists and other Semitists to use them mars many analyses of the prehistoric relations between the Semitic languages and — beyond that — impedes the exploration of more remote connections.¹⁹

Linguistics is, or ought to be, concerned with both advanced research and

elementary pedagogy, to their mutual benefit. Except for a rare individual linguist, however, the two are in fact far apart, and that is to their detriment. The split was perhaps inevitable:²⁰ given the great expansion of linguistic knowledge in the 18th and especially the 19th century, individuals were bound to specialize. Those who endeavor to take cognizance of progress outside of their field, and who apply it pertinently to their own problems, deserve all the more praise. I esteem Konrad Koerner for doing that.

Author's address

Saul Levin
 Department of Classical and Near Eastern Studies
 State University of New York
 Binghamton, NY 13902-6000, USA
 slevin@binghamton.edu

Notes

1. Still earlier, the scholars of Alexandria had perceived the need to supplement the Greek alphabet with accents, etc., which I have treated at length in Levin 1984 and will not summarize here.
2. Eventually appearing in Latin works that deal with Greek as *epsilon*, *ypsilon* (or *upsilon*), *omicron*, and *omega*.
3. These are copiously documented in all but the best manuscripts that survive from antiquity.
4. To be sure, the ancient capital lettering gave way to the medieval minuscule (which we use); but that was a mere circumstance of penmanship, with no effect upon the phonetic function of each of the twenty-four letters: *a* is still the same as *A*, etc.
5. Such a universal negative statement stands as an argument from silence; for even if we could document when so-and-so died, it would scarcely prove that not one person after him pronounced [ai] or [oi]. The total disappearance of a mere phonetic trait, in either a vernacular or a school language, naturally invites less attention than an innovation that offends the ears of conservatives. Even the extinction of an entire language used to take place without being recorded in scholarly literature; but since William Bodener, the last speaker of Cornish, died in 1789, it has been possible to keep track, at least in part, of more and more languages going out of use.

The terms μέγα and μικρόν for Ω and Ο respectively may also have been intended to describe something phonological: the contrasting length of these two vowels, which the grammarians wished to maintain. Or, if it was too late for that, they could still fasten upon the difference in size between the two letters.

6. The ligature Æ, as well as Œ (in minuscule lettering æ and œ), was a sort of compromise between the true ancient digraph AE or OE and the medieval *e* that replaced them. There was an ancient precedent for ligaturing Æ, but only in a word such as the cognomen SABINNÆ

- (*Corpus inscriptionum Latinarum*, I².2.1.639, no. 1602.2) at the end of a line if the stone-cutter, or the scribe, did not have enough space to separate the letters. (Likewise the verb ending NT was often ligatured as in SVN 'they are', 638, no. 1596.2; but the humanists had no phonetic motive, nor any other, for latching on to this ligature.)
7. The humanists did abandon the later medieval minuscule lettering, which they loosely stigmatized as "gothic" (as though it had some real connection with the Goths who had sacked Rome long ago, in 410). They replaced the "gothic" with older minuscule lettering that they called "roman" — and the term remains to this day — but it really developed in France under Charlemagne (around 800) and his successors.
 8. But [y] in France and [ju:] in England.
 9. Milton's Latin poems, I dare say, were all the sweeter because he had the good taste to pronounce the Latin vowels as in Italian, departing from the custom of Cambridge University (where he studied) and the rest of England.
 10. But scorned by men like the hide-bound though lovable Mr. Chipping in James Hilton's novel, *Good-bye, Mr. Chips* (London & Boston 1934), chapter 11. He fastened upon the absurdity of "Kickero" for *Cicero* and ridiculed the restoration of *uicissim* "taking turns" to sound like "We kiss 'im"; he held out for the traditional [vais'sim].
 11. I owe this information to my friend and colleague Mary Ritchie Key.
 12. Much more often हि त म् [hitám].
 13. I prefer *H (= some unspecified laryngeal) to *₂₁, *₂₂, *₂₃ (which many Indo-Europeanists use), because *₂ suggests a semi-vocalic actualization of the weak, unaccentable, nondescript vowel [ə], familiar in French, English, and German: *revue*, *review*, *before*, *bevor*. For I do not know of any language that actually has such a semi-vocalic consonant — let alone three distinct ones.
 14. The unwritten consonant [ʔ] before *æ*-, *eo*-, *i*- has the same alliterative effect as the *s*- in the previous line. See Levin 1992 (250–251).
 15. Cf. also the split vowel of Old English [ʔ] *eorþe* to the alternating {^ε/_ɔ} of Hebrew עָרֶעַץ {ʔérec}, pausal עָרֶץ {ʔārec} (with prefixed article הָעָרֶץ {həʔārec} "the earth"), and הָאָרֶץ {ʔārac^h}, pausal הָאָרֶץ {ʔārac^h}: Greek ἑρᾶς "earthward".
 16. This is preferable to the simple but misleading transcription *a*, *i*, *u* of some Semitists, which unfortunately suggests a pure vowel.
 17. Drastically reshaped to *l* in Arabic.
 18. These are termed *cultismos*.
 19. E.g., in the admirably comprehensive book by Orel & Stolbova 1995.147, under "639 *dam- 'blood'" the Hebrew דָּם — which is almost invariably accented throughout the Bible — is cited in the misleading conventional transcription "Hbr *dām*", rather than {dóm}. This obscures the resemblance of the word in one Semitic language to the correspondents in several West Chadic languages: "WCh *dam- 'blood': Sura *tɔɔm*, Ang[as] *tom*, ... Bol[ewa] *dom*, T[a]ng[ale] *tɔmm*, N[a]gm[a] *dom*, Maha *dom*, Bele *dom*, ... Gera *n-dooma*". A further consequence is the biased reconstruction of the West Chadic proto-form *dam- (to match "Sem *dam-"), with a vowel not found in any West Chadic language except "K[i]fr[i] *n-daame*", rather than the back vowel that actually prevails in this group of languages. See Levin (1989: 98–99).
 20. Harris (1987: 108–115) singles out Ferdinand de Saussure as a major influence, determined to make linguistics a respectable, independent science.

References

- Corssen, Wilhelm. 1858–59. *Über Aussprache, Vokalismus und Betonung der lateinischen Sprache*. 2 vols. Leipzig: B.G. Teubner.
- Harris, Roy. 1987. *The Language Machine*. Ithaca, NY: Cornell University Press.
- Levin, Saul. 1984. "The Greek Diacritical Marks and Their Application to Other Languages in the Renaissance". *General Linguistics* 22.21–37.
- . 1989. Review of *Spelling in the Hebrew Bible*, by Francis I. Andersen & A. Dean Forbes (Rome: Biblical Institute Press, 1986). *Hebrew Studies* 30.96–100.
- . 1992. "Semitic Evidence on Some Problems of Indo-European Prehistory". *Word* 43.251–265.
- Orel, Vladimir E., & Olga V. Stolbova. 1995. *Hamito-Semitic Etymological Dictionary: Materials for a reconstruction*. (= *Handbuch der Orientalistik*, 1. Abt., 18. Band). Leiden: E. J. Brill.

CHAPTER 23

The Autonomy of Linguistics: **Saussure to Chomsky and beyond**

Ranko Bugarski
University of Belgrade

1. Introduction

The question of the autonomy of linguistics relates to the issue of the emergence and survival of a separate science of language with its own object of investigation.¹ It must therefore be studied in the interaction of these two aspects — the linguistic, concerned with the object, and the metalinguistic, having to do with the science. Throughout its history modern linguistics has been crucially marked by an effort to lay the conceptual and institutional foundations for a science of language which would no longer be subordinated to philology, philosophy, sociology, or some other larger discipline, but rather a self-contained and legitimate member of the circle of modern sciences. And the way to achieve this goal was sought in an attempt to define language as a distinct system, irreducible to subsets of its manifestations which could be individually and adequately covered by the other sciences just referred to.

In other words, as opposed to disciplines whose areas of interest include some aspect or another of language as well, and which therefore tend to see in language a means of gaining knowledge of extralinguistic phenomena of some kind, linguistics came to perceive in it a coherent whole and a worthy object of concentrated scholarly study. Such a conception presupposed the existence of a linguistic system which could in principle be described independently of the history, culture, social context, or individual peculiarities of its users, and in that sense it may properly be identified as the postulate of the autonomy of language and linguistics.

2. The autonomist position

Elements of such a view of the immanent structure of language had occasionally been in evidence in the history of language study; particularly relevant from this standpoint was, for example, Neogrammarian insistence on the mechanical laws of sound change. But this position was elevated to the rank of a comprehensive program only with the establishment of contemporary linguistic theory, in the form of a call for the abstract investigation of the synchronic structure of linguistic systems, founded on the postulate of the primacy of form over substance in language. Accordingly, the label of autonomous linguistics can with full justification be applied to structuralist orientations of the period between Saussure and Chomsky, the same one in which modern linguistics came of age and achieved its greatest theoretical and methodological results.

It is precisely in the context of the quest for autonomy that one should reflect on the postulation of abstract and idealized objects like 'langue' or 'competence', with an emphasis on the structure of grammar. The more superficial layers of language structure, and particularly the actual manifestation of language systems in the process of speech, were regarded as too much conditioned by physiological, neuropsychological, social, contextual, and other extralinguistic factors to warrant isolation from the speech chain and exact scientific exploration.

After Saussure's sharp and deep cuts had laid firm foundations for an autonomous linguistics, his followers proceeded along the same lines. Thus Hjelmslev affirms the view that language is not a mere collection of physical, physiological, psychological, or logical phenomena amenable to piecemeal research, but a system *sui generis* with its own irreducible essence:

A linguistic theory which searches for the specific structure of language through an exclusively formal system of premisses must ... seek a *constancy*, which is not anchored in some 'reality' outside language — a constancy that makes a language a language ... When this constancy has been found and described, it may then be projected on the 'reality' outside language, of whatever sort that 'reality' may be ... so that, even in the consideration of that 'reality', language as the central point of reference remains the chief object — and not as a conglomerate, but as an organized totality with linguistic structure as the dominating principle (Hjelmslev 1961:8).

He concludes his treatise on the theory of language with a vision of "immanence and transcendence ... joined in a higher unity on the basis of immanence" (Hjelmslev 1961:127).

This autonomist *credo* of the Danish theoretician from 1943, including the

conclusion which does not preclude studying the external side of language but provides for it only on the basis of prior knowledge of its internal organization, symbolizes in its own way a fundamental aim and orientation of modern linguistics. This particularly applies to those central and by far most influential versions of it which can most justifiably be described as structuralist. And by this we mean not only the European schools but also American structuralism, which from this point of view passed through essentially the same phases of development. Here the principal role in the establishment of linguistics as a science fell to Bloomfield and his followers (see especially Bloomfield 1926, 1933), while the appearance of Chomsky on the scene was in some sense to mark the logical culmination of the continued striving for autonomy (though in his case of language rather than of linguistics; see below).

In summary fashion, we might attempt a joint reconstruction of European and American developments — to be sure, only in a rough and schematic manner — in terms of three main and overlapping stages of increasing autonomy. First, independent study of language was affirmed through the consolidation of linguistics, as against all the other sciences partially involving linguistic subject matter. This study was then narrowed down to the synchronic investigation of language structure itself — that is to say, divorced from the context of history, culture, and society. Finally, the focus of the notion of language structure was laid in the domain of grammar in the strict sense, notably syntax as its most abstract and to that extent most autonomous section (in other words, the 'covert' structure of deep relations, as opposed to the 'overt' morphophonological structure).

Striving for autonomy, in the case of linguistics as of other sciences in their formative phases, is natural enough. Yet it has constantly met with opposition too. Thus, for example, Schuchardt, Gilliéron, and others spoke out against Neogrammarian mechanistic positivism, orthodox structuralism came under attack from different quarters, and resistance to generative grammar has likewise not abated from its very inception. What is common to all these voices of dissent is that they have tended to perceive in autonomous linguistics, with more or less justification, an artificial and actually harmful severance of language from its natural historical, social, and communicative context. Yet reasonable as such objections may be in themselves, it needs to be said that in its own day autonomy was the only possible programmatic option, without which linguistics as we know it today would never have come into being.

3. The meaning of autonomy

A different and crucially important matter is the changing sense of autonomy. After autonomous status had been achieved to a significant degree — in some aspects by about 1930, in others a decade or two later — it was no longer necessary to keep stressing its importance. From that time on linguistic autonomy is at the same time taken for granted and relativized. In addition to activities of the predominantly ‘heteronomous’ orientations of anthropological, sociological, and similar schools of thought (e.g., Boas, Sapir, Malinowski, Firth, etc.), functionalist directions within structuralism itself (especially the Prague School as represented by Jakobson) considerably modify extreme autonomist positions, announcing an era in which it would be possible, from a secured autonomous foundation, to recognize new and wider horizons. And a decade or so later the question of the autonomy of language and linguistics was to be raised anew, though in a different way, by the establishment and fast development of the various interdisciplinary areas of language study (such as psycholinguistics, sociolinguistics, neurolinguistics, pragmatics, applied linguistics, later still cognitive linguistics, etc.).

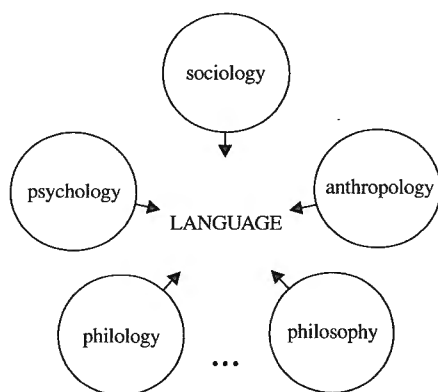


Figure 1. Heteronomous (pre-structuralist) stage

The course of events we have just sketched out can be represented in a simplified graphic form, in three stages. The first may be roughly labelled ‘heteronomous’ or ‘pre-structuralist’ (Figure 1). The second is ‘autonomous’ or ‘structuralist’ (Figure 2). The third is by and large the present one, which we



Figure 2. Autonomous (structuralist) stage

shall identify as 'autonomous-interdisciplinary', having in mind a rough label of the type 'enriched structuralism' (Figure 3). (In this sequence the label 'post-structuralist' would come in handy, but we avoid it as not really capturing the essence of the present situation of linguistics, whose central domains — at least as this writer sees them — are still dominated by an extended and updated version of structuralism; see Bugarski 1989 for details of this view).

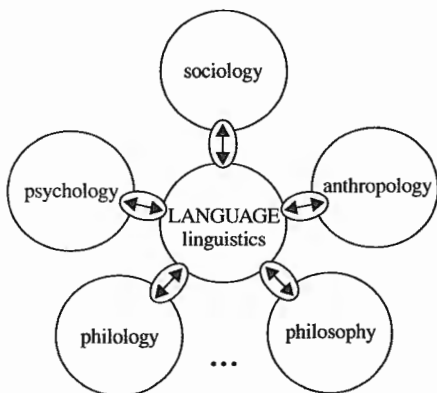


Figure 3. Autonomous-interdisciplinary stage

What can be said, then, about the present-day linguistic scene from the point of view of autonomy? This is no occasion for a detailed examination, so we can only point to those aspects of the situation which seem most important. First of all, and most generally, the contemporary growth of interdisciplinary and applied fields of research within the broad domain of linguistic phenomena has at least in some measure called into question the appropriateness of further discourse on autonomy, at least in the earlier framework of a rather radical understanding of this notion. That is to say, if cardinal aspects of language such as the psychological, the sociological, and others are severed from linguistics and brought under the jurisdiction of specialized disciplines like psycholinguistics or sociolinguistics, it would appear that linguistics itself would be left with a considerably reduced autonomous object and with it, presumably, status as well. One should also mention here considerations of autonomy in the context of debates on the epistemology of linguistics, involving attempts to determine its standing as an empirical or, partly, axiomatic discipline. Thus in accordance with one view, which has stimulated further critical and polemical discussion, linguistics under its central grammatical aspect is a non-causal, essentially non-empirical, and hence autonomous discipline, while in its interdisciplinary portions, notably psycholinguistic and sociolinguistic, being fundamentally causal and empirical, and so non-autonomous. (See Itkonen 1978, 1984 with references; also, for example, the polemical exchange between Kac 1980 and Prideaux 1980).

However, an effective reduction of the proper domain of linguistics along the lines just alluded to must from a present-day vantage point be regarded as unwarranted and harmful. In the totality of its range linguistics legitimately encompasses the major interdisciplinary branches such as psycholinguistics and sociolinguistics — and not as marginal segments but as central areas of a general science of language. (For argumentation supporting this position see Bugarski 1987). And as far as autonomy is concerned, it would appear that the real source of many misunderstandings and fruitless debates is in the varying and often quite inadequate interpretations of this concept. For if in the formative years of linguistic structuralism, autonomy could be and probably had to be conceived rather strictly, at least at the programmatic level, and not infrequently with a high degree of exclusivity, the times are different now. Much has already been achieved, the science of language has been stabilized, and as a result of this conceptions of the meaning and range of autonomy are of necessity fundamentally modified.

4. From independence to integration

It appears that the most serious mistake is made when the notion of autonomy, even if only implicitly, is taken in an absolute sense and identified with that of independence. Yet these two concepts are not synonymous. Autonomy means self-reliance, without however connoting self-sufficiency — and it is precisely this connotation which is often linked with independence. If independence leads to such freedom of behavior and action as permits no 'interference' of external motives, criteria or models, whereby any broader perspectives are blocked by arbitrariness and ultimate futility, we are still further removed from the meaning of autonomy. Whereas independence may mean not only the absence of dependence on other fields but also the lack of any links whatever with them, autonomy by no means presupposes the latter. Such links may even acquire the character of interdependence, without involving a violation of autonomy (the basic prerequisite for autonomy being, of course, the absence of unilateral dependence — that is, of actual subordination). Hence mutual dependence of autonomous disciplines as equal partners, even to the extent of partial integration into a larger whole, is in no way a self-contradictory concept.

These relationships also explain the apparent paradox whereby linguistics, developing in recent times as an autonomous science, has in so doing strengthened rather than weakened its ties with psychology, sociology, anthropology, logic, and a number of other fields. One might even say that in the process there has occurred a noticeable reduction in the autonomy of its very object, caught up in the numerous and multicolored spotlights of all these disciplines, but this has not been accompanied by a corresponding reduction of the autonomous status of linguistics. Simply, interdependence does not entail the loss of autonomy and a return to an earlier condition of unilateral dependence. Whatever the extent of its cooperation with other sciences, including the possibility of its integration with them in the future, linguistics retains its autonomy in that it no longer is — nor is ever again likely to be — merely a sidetrack of some other science.

This brings us to the question of the relationship of linguistics to some hierarchically superordinate science whose part it might constitute. In freeing itself from an earlier dependence on other dominant disciplines and in achieving autonomous status, linguistics removed this question from its agenda in an operational sense, but it still presented itself from time to time on the level of general principles and broad perspectives. Let us not forget that in this sense Saussure himself saw linguistics as a branch of semiology and hence of social psychology; yet this did not prevent his work from being built into the very foundations of autonomous linguistics.

Similarly, Chomsky regards linguistics as essentially a division of cognitive psychology, without however implying a loss of its autonomy. This would occur only if nonlinguistic fields of specialization were to dictate to linguistics what it should do in its own house, proclaiming its legitimate goals, formulating relevant problems for it to investigate, forcing on it research methodologies, etc. Yet nothing of the sort is advocated by Chomsky or by any other serious author in linguistics. For Chomsky himself this is, then, apparently not an issue of linguistics being dependent on cognitive psychology, but merely one of the general orientation of linguistic theory as a discipline which may contribute to revealing the secrets of the human mind and of the nature of our cognitive capacities. (Chomsky in fact reaffirms an earlier tradition of equal partnership in a joint venture when he describes the insistence of linguistics, psychology, and philosophy on "emancipating themselves" from mutual "contamination" as a "peculiarly modern phenomenon"; see Chomsky 1966: 76). At the same time it is foreseen, naturally enough, that linguistics itself will benefit from the achievements of other sciences, as part of the normal interaction and cooperation of scholarly disciplines in the present-day world.

Very briefly put, singling out the individual branches of general linguistics, associated with the cardinal aspects of language, could place the science of language within the scope of different overarching fields of study: of semiotics if "semiolinguistics" (the investigation of language as a code or sign system) were considered the most important, of psychology if the psycholinguistic component were to be favored, and of sociology should precedence be given to the sociolinguistic segment. All these theoretical positions, which would in effect accommodate linguistics as a whole under the roof of some other science, have been authoritatively advocated — or at least implied — in the recent development of the science of language. In our own view, however, such a reduction of linguistics to any one of its principal branches is to be rejected. Likewise, although language is the main means of communication among humans, construing linguistics as part of a general science of human communication would represent a largely artificial solution, since such a discipline hardly exists at present. In the process of its possible constitution linguistics itself would necessarily claim a central position, but as the notion of communication is considerably broader than that of language, the domains of these disciplines could scarcely be seen as overlapping, even in theory. Hence it does not seem realistic to regard linguistics as a branch of the science of communication, or to declare linguistics itself to be such a science.

If one still insisted on identifying a more comprehensive discipline, the basic fact that language is a property of human beings would presumably entail

that the proper perspective might be to treat linguistics as part of a larger science of man. In terms of our current nomenclature such a discipline could best be called anthropology, but in a new and broader sense yet to be elaborated in a combination of the contemporary scientific and philosophical meanings of this term. Therefore this solution too is merely conditional and, for the present at least, theoretical much more than operational.

As we have already pointed out, it is essential for this discussion to note that cooperation with other sciences, and even partial integration with them, does not really mean the loss of autonomy. On the contrary, autonomy if flexibly conceived shows itself to be a relative concept which should in fact be linked with integration rather than with independence. Interpreted as absolute independence, autonomy leads only to the deadlock of isolation, but it gains its proper perspective when regarded as complementary to integration — naturally, under an equally flexible interpretation of the latter concept. Such a warning was sounded by the far-sighted Roman Jakobson in his review of the relationship between linguistics and other sciences at the Bucharest congress of linguists in 1967 (see Jakobson 1969: 75–76).

It appears that the relevance of this warning has increased in proportion to the further development of such relations during the last three decades. Throughout this period we have heard influential spokesmen for the independence of newly created subdisciplines or interdisciplines, but in assessing such pronouncements it is important to differentiate purely scholarly argumentation from the apparently dominant motives of a pragmatic character. In other words, early independence may well be desirable from the point of view of individuals or institutions involved in these new fields. But if we regard the latter primarily in their scholarly aspects, from the standpoint of building up theories and methodologies appropriate to their subject matter and objectives, such fragmentation can hardly be beneficial in the long run. Furthermore, it is difficult to resist the impression that, given the normal changes in living conditions and in the technology of scientific research procedures, what has been inherited as indispensable intellectual autonomy may at times be imperceptibly transformed into profitable institutional autonomy, the real purpose of the latter possibly even receding from view in the process.

5. Conclusion

In the sense in which we have used the term, autonomy remains a feature of present-day linguistics and, in a significantly decreased measure, of language as

its object too. In the science of language, autonomy manifests itself in varying degrees. The highest degree still attaches to linguistics in the narrow sense, namely the study of language structure itself, but interdisciplinary fields like psycholinguistics and sociolinguistics, with their feedback effect on that core area, exhibit a certain measure of autonomy as well. Knowledge derived within them is integrated into linguistics as a whole. Analogously, this entire science finds itself in growing interaction with numerous other disciplines, tending towards integration on equal terms into the system of sciences — an integration which does not mark the end of its autonomy. (For some remarks on integration in this context see Bugarski 1990). The whole question of autonomy is thus only a relative matter, since in science as in life nothing really exists in and of itself, apart from all the rest — even though individual phenomena must in the nature of things often be singled out for separate study or experience.

Author's address

Ranko Bugarski
Dalmatinska 11
YU-11120 Belgrade 35

Notes

1. It is a pleasure to offer this small contribution to Konrad Koerner, respected scholar and friend with whom I have had the privilege of collaborating over many years.

References

- Bloomfield, Leonard. 1926. "A Set of Postulates for the Science of Language". *Language* 2.153–164. (Reprinted in *Readings in Linguistics* ed. by Martin Joos, 26–31. New York: American Council of Learned Societies, 1957.)
- 1933. *Language*. New York: Henry Holt.
- Bugarski, Ranko. 1987. "Applied Linguistics as Linguistics Applied". *The Relation of Theoretical and Applied Linguistics* ed. by Olga Mišeska Tomić & Roger W. Shuy, 3–19. New York & London: Plenum Press.
- 1989. "Generative Structuralism". *Yugoslav General Linguistics* ed. by Milorad Radovanović, 33–46. Amsterdam & Philadelphia: John Benjamins.

- 1990. "Towards Integration in Linguistics". *Proceedings of the Fourteenth International Congress of Linguists* ed. by Werner Bahner, Joachim Schildt & Dieter Viehweger, vol. III, 2612–2614. Berlin: Akademie-Verlag.
- Chomsky, Noam. 1966. *Cartesian Linguistics: A chapter in the history of rationalist thought*. New York: Harper & Row.
- Hjelmslev, Louis. 1961. *Prolegomena to a Theory of Language*. Translated by Francis J. Whitfield. Revised English edition. Madison, Wisconsin: The University of Wisconsin Press.
- Itkonen, Esa. 1978. *Grammatical Theory and Metascience*. Amsterdam: John Benjamins.
- 1984. *Causality in Linguistic Theory*. London: Croom Helm.
- Jakobson, Roman. 1969. "Linguistics in Its Relation to Other Sciences". *Actes du Xe Congrès International des Linguistes* ed. by Alexandru Graur, vol. I, 75–111. Bucarest: Académie de la République Socialiste de Roumanie.
- Kac, Michael B. 1980. "In Defense of Autonomous Linguistics". *Lingua* 50. 3.243–245.
- Prideaux, Gary D. 1980. "In Rejection of Autonomous Linguistics". *Lingua* 50. 3.245–247.

CHAPTER 24

On the Notion of Zero

E. Wyn Roberts
Simon Fraser University

1. Introduction

In the history of linguistic theoropoesis specifically, no less than in the history of theoropoesis in general, far too little attention has been directed towards the notion of zero. While clearly present from early times as a digit — and hence place-value — in, for instance, Hindu and Mayan number systems, its appearance in Western number systems was late — and even then not an indigenous development, being introduced through the adoption in the West of the Hindu-Arabic number system. As far as recorded history reveals, the concept was first introduced into linguistics by Pāṇini.

Several problems have plagued its proper conceptualization. Putatively synonymous or paraphrastic definitions of zero in English are ambiguous to the point of compound confusion. Consider zero ‘defined’ as the following: ‘nothing’ (or ‘naught’); ‘not any’; ‘none at all’; ‘the complete *absence* of quantity’; ‘the *absence* of something’; ‘the *absence* of everything/anything’ (all of which circuitously imply the equation of zero with nothing, nothingness, absence) vs. ‘the *presence* of nothing’, which implies the equation of zero with *presence*, and thus, with something, some value. It is also occasionally defined as ‘the lowest point’, and, for the expression ‘zero hour’, as ‘the time set for the beginning of an attack’. Analogously, compare the definition of absolute zero temperature (relatable to the definition of terms such as chaos, randomness, entropy, etc.) as the temperature at which a system is apparently devoid of all energy, and so of all motion — in the sense that molecules at least appear not to oscillate at this stage. The terms ‘nothing’, ‘presence’, ‘absence’ are, of course, themselves equally problematic from a definitional point of view, as is the notion of

definitionalism itself. Clearly the likelihood of infinite regress in this discursive approach is very obvious.

Consider an English expression such as "I have no money (in my pocket)". This may either be an answer to the question "How much money do you have (in your pocket)?" or followed by the question in the following form: "HOW much money do you have (in your pocket)?" In this case, 'no' is a quantifier, marking zero money (in the pocket). That is to say, there is money in the pocket, but its quantity is zero. If this were not the case, then neither the question nor the answer would make any sense. The expression "I have nothing (in my pocket)" also states that there is some quantity zero in the pocket. It is not the same as saying that the pocket is completely void and unstructured. This is zero as a quantifier. In mathematics, zero as place value (in linear order) is as near to empty or 'void', or 'vacuum' as possible. However, there is a value to that place, or void, or vacuum, of, say, relatively least possible energy or matter. This is zero which is physical and material.

The postulation or establishment of zero elements has been invariably treated within the assumption of the phonetics:linguistics dualism as part of phonological rather than of phonetic theory. Thus, the notion of zero is inextricably interwoven into the fabric of dualisms that characterizes theoropoesis not just in an Aristotelian epistemology but globally. Thus, the form:meaning, phonetics:phonology, phonology:grammar dualisms, as well as the principle of the arbitrariness of the linguistic sign, all have direct bearing on the attitude towards zero. It is not surprising, therefore, that, without exception, the notion of zero (\emptyset) has caused a great deal of difficulty in every school of linguistic thought, and the following statement by Haas (1954: 77), made almost half a century ago, is no less true today than it was then:

We may not have succeeded yet in clearly stating the conditions in which it is legitimate to speak of the presence of an element *zero*.

At about the same time, Allen (1955) could comment:

In reading contemporary linguistic statements one encounters the term 'zero' in so many and diverse contexts that one may be forgiven for wondering whether its use may not be exceeded by its abuse.

Allen (1962: 212) adds that:

the linguistic device of 'zero' is justified only by the requirements of general statement; in order to extend the application of a given grammatical structure, the system established at a place in that structure is made to include a term for which, unlike the other terms, no phonological bit is statable.

and

in no particular case is it essential to employ the concept at all, and that its use arises *only from the requirements of generalization*. [My emphasis, WR]

However, since linguistic theory, like all scientific theory, is about the *necessity* of (maximal) generalization, to which particular cases are subordinate, the 'use' of any concept that can contribute to (maximal) generalization is in the final analysis not just a matter of choice. On the contrary, its theoretic adoption is *essential*, and any theoropoesis that to an ad hoc degree falls short of the requirement of (maximal) generalization is clearly self-contradictory since it constitutes a negation of the first and inevitable principle of generalization. For the same reasons, any apriori arbitrary constraints on any such notion is unacceptable.

All linguists who have seriously considered using some notion of zero in linguistic statements have found it necessary to append *caveats* of the most general kind that are more moralistic than rational. Thus Harris (1951:339, fn.29) warns that:

the indiscriminate use of zero segments and void elements can make many different language structures seem sterilely similar.

Such arguments are patently circuitous since they are based on unknowns — such as how to determine inter-language similarity/sameness — in other words how to conceive of the distinction between universal (i.e., linguistic theoretic) and typological (i.e., statistical) statements. This is a problem of what Conant (1957) has aptly termed the degree of empiricism of one's statements.

Similarly, Nida (1947: 46) states that:

the description of a language becomes unduly sprinkled with zeros merely for the sake of structural congruence and balance.

Even if linguistics is innocently characterized as the systematic study of language and the study of language systematicity, clearly, then, the establishment of what constitutes linguistic 'structural congruence and balance', and thus 'similarity' between languages, is the prime theoretic concern in linguistics. Therefore, it is crucial to understand just as clearly what the criteria are for 'indiscriminate use [and 'undue sprinkling'] of zero segments and void elements', 'for 'zero' vs. 'void' elements', for 'seeming' similarity between 'different' languages, for 'similarity' vs. 'sameness' between languages, for 'mere' or 'just' vs. self-evident and justifiable, and for 'structural congruence and balance' vs. whatever the contrary might be conceived to be. Simple authoritarian fiduciary *caveats* must not be allowed to replace principles, and conceptual pusillanimousness cannot displace demonstrated logic, and reductionism *sensu bono*. A decision to

use or not use zero requires a complete commitment to a true and honest evaluation of the ultimate consequences of that decision. Aprioristic foreclosures are as immoral as they are unscientific, as the history of our own and of every other discipline has shown.

2. Brief overview of the use of zero in linguistics

Saussure (1915 [1955]: 123–124) states for modern linguistics the basic hypothesis concerning zero thus:

On voit donc qu'un signe matériel n'est pas nécessaire pour exprimer une idée; la langue peut se contenter de l'opposition de quelque chose avec rien.

The principle enunciated is here, as generally, exemplified in morphology, especially affixal morphology. This reduces to saying that 'zero' is a semantic (correlation) notion. The use of zero in phonology is apparently less attractive, or 'more difficult'. The reasons for this — although attractiveness and difficulty have no relevance in theoropoesis — should be reasonably clear from the presence in Saussure's statement of the words '*matériel*' and '*rien*'. An '*idée*' is *not* material, and as such is not necessarily materially expressed. This argument can be seen to be directly related to, or derived from, the principle of the arbitrariness of the linguistic sign, of the relation between meaning and sound.

American Structuralists also appear to have had no problem with the use of zero in morphology (cf. Bloomfield 1933: 209). However, even at this level, it was more commonly the case to have a zero morph ('allomorph') than a zero morpheme, i.e., a morpheme where there was only one allomorph — zero — or where all the allomorphs were zero morphs (consistent with the contextual, complementary distribution criterion of emic identity). This kind of zero at the morphological level is a more generally acceptable concept than is zero phoneme or zero allophone at the phonological level. Nor was there any problem with the use of the device in morphophonemic alternation (Harris 1951: Chap. 14; Bloomfield 1933; see below on transformational generative phonology [TGP]), a traditional concept clearly necessary in diachronic linguistics as well. The fact of homonymy as part of the role of 'meaning', in whatever sense, in linguistic analysis, regardless of where it became apparent in the methodology, made the notion of overlapping allomorphs a necessary one. Consequently, it is impossible to impose upon morphology the concepts of invariance and biuniqueness, and once the concept of overlapping of allomorphs is allowed, any conditions disallowing zero morphs/allomorphs similarly break down.

Both the Bloomfieldians and the Neo-Bloomfieldians accept the use of zero in morphology (Harris 1951; Hill 1958; Gleason 1961; Nida 1947). The only constraints are general ones, as in Harris (1951: 334):

..., this procedure is never unavoidable: in any corpus of material, it is possible to identify every linguistic element safely in terms of non-zero stretches of speech.

with which we may compare Gleason (1961: 76):

We can never add zeros *beyond the limits of the gaps clearly visible* in the structure being described.

Robins (1964: 204), as representative of the Firthian brand of British structuralism, further remarks:

ZERO elements, the name given to analytic entities *without actual or overt exponents*; [they are] justified if the introduction of the zero element *simplifies the resulting analyses by aligning the structures of more words together*. [My emphasis, WR]

This statement must, however, be interpreted in relation to phonology in the context of Firthian polysystemicness (see, e.g., Robins 1957), wherein the concepts of phoneme and allophone are irrelevant. It is clear that the constraints and allowances cited are invariably extremely vague, or their logical implications are not followed through upon.

In phonology, however, the attitude towards zero is, to say the least, a much more stringent one. Bloomfield (1926: 27) disallows 'zero' as allophone or phoneme although it functions within morphophonemic (1926: 29a) and morphemic (1933: 296) statements. Swadesh (1934: 37), for whom morphemes consist of phonemes, allows neither zero phoneme nor zero allophone (although there must be *deletion*, i.e., a zeroing or a voiding processes in morphophonology). However, in the works of the early structuralists, there are many references to zero which imply either that zero phonemes are allowable or that the general linguistic foreclosures on its use or non-use had not yet been considered. There are also clear indications that the attitude towards zero may often have been vaguely typological, even political, tolerated for 'exotic' languages such as Chinese, but not for 'non-exotic' languages such as English, German, or French. Chao (1934: 43a), for instance, makes the following initial statement:

(c) **Zero Symbols** *As limiting cases of the variation in size of unit* [My emphasis, WR], we have the possibility of using zero symbol for sounds or sound-features and of counting absence of sound as a phoneme or as one member of a phoneme.

Twaddell (1935) apparently finds no use for the concept, and accords it no mention. Bloch & Trager (1942: 40) similarly espouse a strongly anti-zero approach:

The process of discovering the phonemes of a language is essentially one of arranging, comparing and combining the formal utterances (and parts of utterances) recorded in a *phonetic transcription*. [My emphasis, WR]

Thus, *what is not recorded* (i.e., 'zeroes') never enters into consideration in phonemic analysis, and no question of zero phoneme or zero allophone arises. Here it is crucial to note that the problem of zero is always couched in *formal* terms — i.e., whether it is a term at some *abstract* linguistic level, or interlevel. There is no mention of *phonetic* zero. The putative *concreteness* of phonetics seems to make zero a non-issue, since it is at that level of interpretation of speech behavior (the *phonic* material according to Firth) that it is invariably tacitly assumed that the decision can and is made about whether there is some speech element 'present' or not. It is assumed to be an easy and self-evident process! Yet, clearly, it is not. Phonetic representation is no less an abstraction that is phonological, morphological, or syntactic. In fact, the only thing that in principle distinguishes, say, phonetic from phonological representation is that the latter is a *functional* representation *matching semantic and sound differentials*. It is in this sense that phonology is semophonetic, functional, correlational, and, thus, at least bi-disciplinary. It is this that also makes the two types of statements incompatible. There is no reason why what is generally regarded as an utterance beginning with a member of the 'general phonetic' class vowel should not be interpreted phonetically as a sequence of a smooth 'consonantal' transition (zero, 'glottal' glide/approximant) followed by a glide, or liquid, or vocalic element. This is in fact well evidenced in orthographies which are more interesting and imaginative than the roman — e.g., Greek, Korean Hangŭl, Semitic. It is also clear that the redundantism assumed necessary to establish meaning:sound differentials, i.e., phonemes, is then applied equally to phonetics, in which meaning differential should play no role whatsoever. Thus, in the present paradigm of linguistics, the relation between phonetics and linguistics is viciously circuitous by virtue of the pervasiveness of 'meaningfulness' as the prime consideration.

It is interesting to note Joos' (1957: 123a) comment on Hartman's (1944) article:

Now to me at least it seems obvious that whatever is completely predictable and has no distinctive features of its own is no phoneme, whatever else it may be. I can find no better name for it than 'zero'.

which links 'zero' and predictability. However, in statements about speech-language which are not driven by some vague semantic notions, there is clearly missing a theory in which so-called contextual or predictable features are universally explained. This being the case, such predictabilities — including the notion of complementary distribution — are merely restatements of observations, and are in no way an exegesis. The relation between 'phonetic' and 'phonological' statements is discussed further below in the context of considering the Naturalness Condition.

The criteria of invariance and biuniqueness, in the narrowest sense, clearly have a strong hold over phonemics. There was nothing comparable to morphemic homonymy in phonology which would allow overlapping of allophones, and apparently no motivation to generalize structures beyond the syllable (essentially the largest meaningless *form*). As will be shown later, canonical syllables — which are in fact more *formulaic* than formal — fulfill one function and one function only, namely to indicate the fullest extent of distributional statements for the terms constituting the symbology of the canonical forms themselves. The form of the 'canon' was to fulfill no further role in defining, redefining, or reducing the elements comprised in its statement. Both the elements and the canonical statements are final-order ones.

The issue of 'zero' in post-Bloomfieldian phonemics is also closely related to the notion of *juncture*. Hockett (1947: 221 et passim), in his treatment of Peiping phonology, like Hartman, finds it necessary to make extensive use of the concept 'zero' (marked #) without, however, entering into any scientific discussion of the matter. Yet, in 1966, A. A. Hill could still say that:

It is a serious criticism of someone working in strict phonemic terms if he sets up a phoneme which can contain zero as a member. If the analyst is describing morphophonemes, zero members are legitimate enough.

although it is not clear whether this allows a zero morphophoneme or 'just' a zero morpheme/allomorph. The functional nature of junctural phenomena had been known for some time but it is in the work of Bloch & Trager (1942) that its relevance for phonemic analysis becomes a central concern, beyond the notion of pause as related to secondary phonemes of stress and pitch as Bloomfield (1933: 114 sqq.) conceived it. The general issue of juncture in phonemics is not of prime concern here, except insofar as it relates to a functional zero concept. Basically, there are two conceptions of the notion of juncture, associated generally and respectively with Hockett and Harris. Hockett's concept (1955: 167–172) was that:

junctures are in some ways a very different kind of phonologic entity from ordinary vowel or consonant phonemes

in that the heuristic principle of 'phonetic realism'

must be not merely suspended but almost reversed. (1955: 158).

Compare Robins (1964: 176):

The juncture phoneme has as its members all the phonetic differences abstracted from the sequences involved which by the formal method were assigned to the neighbouring segmental phonemes. [Cf. Jakobson et al. 1951: 43 (DF matrix for English phonemes), and 1951:39. The lax counterpart of /h/ presents an optional variant: in cases of emphasis a glottal catch may be substituted for the even onset: *an aim* can appear in the form [ənʔ'eim] in order to be clearly distinguished from *a name* [ən'eim].]

In this sense, then, there is no zero phoneme at all, rather a phoneme whose allophones are 'more varied' than the 'standard' types (this statement clearly approximates the nonsensical since, in principle a phoneme has [at least] as many allophones as there are contexts in which it appears), and whose establishment requires a variation in procedure. For Harris (1951: Chap. 8.), however,

the juncture phoneme is itself a zero element (a unit required and justified by overt observable phenomena, but not itself comprising any variants or allophones).

This is a most startling statement — a phoneme with no allophones! Yet, this zero phoneme provides the different phonological environment by which the phonetic features involved in the syllabification /ən'eim/, /ət'i:z/, etc., occur as allophonic variants (Robins loc.cit. and also Moulton 1947). Within phonemics, the different concepts of 'juncture' involved crucial implications. Hockett's approach clearly maintains standard phonemic percepts and, hence, no zero phoneme and no zero allophone is effectively involved. In Harris, however, there is, and it is this notion that has drawn the greatest amount of criticism. Note, however, that the purpose in each case is exactly the same, viz. to minimize the number of consonantal and vocalic phonemes for any language. There is no consonantal zero phoneme (whatever this may mean), and no vocalic zero phoneme (again, whatever this may mean; cf. Chao 1934; and Joos 1957: 123).

Within the Item and Arrangement approach morphemes are made up of phonemes, and thus it follows that it is impossible to conceive of a zero morphophoneme. Zero functions only in situations where a phoneme in one allomorph corresponds to nothing (i.e., a zero) in another allomorph. However, since allomorphs do not contrast, there is no contrast (morpho)phoneme vs. zero. This view, viz. that a morphophoneme cannot be zero, only voidable, is adopted into Item and Process models as well, including TGP (see further below). Undoubtedly, the paradigm established in historical phonology contributed to the attitude to

zero in both schools. In American structuralist phonemics, as in fact generally, the deciding factor appears to have been the strong degree of semanticism (implying the meaninglessness of whatever is interpreted as 'nothing') — a form of high level of empiricism — directing a strong form of invariance/biuniqueness condition on the statement of phonological contrasts.

2.1 *Praguan phonology and zero*

The definition of the phoneme and the rules given for determining phonemes clearly indicate that Trubetzkoy would not consider establishing a zero phoneme (Trubetzkoy 1939 [1969]: 27). This view is further supported by the definition of *opposition* (1939: 31). It is interesting, moreover, that so-called pre-vocalic glottal stop, for instance in morpheme-/word-initial position in German, is not treated by Trubetzkoy as phonemic but as a feature of the vowel — "It is not a phoneme but merely a 'natural way of pronouncing' vowels in morpheme-initial position". However, a glottal stop that contrasts with smooth off/on glide (as in word final post-vocalic position in Finnish) is phonemic (*ibid.*). This apparent anomaly in treatment is understandable only in light of the priority of system economy. In German, if we accept [ʔ] as phonemic, all we do within the available phonologies is add a unit to the system, a unit which can in fact be predicted in terms of otherwise necessary elements. A statement to the effect that, given /ʔ/ in German, all words (morphemes) begin with a consonant *is not worth the cost of an extra phoneme*. Extension of the consequences of postulating /ʔ/ as of postulating /Ø/ appears never to have been considered. In phonemic analysis, doing either (or both) would merely have effected an increase in the *system* and, hence, a lack of economy. Whether this particular type of economy of system over structure (an issue only superficially considered in relation to the analysis of affricates and, occasionally, diphthongs) was merely an immediate rather than a long-term one if further re-analysis were undertaken, was never seriously considered. Indeed, even the one linguist to take the concept of re-analysis seriously, viz. Harris, failed to attain any lasting effect on linguistic theory.

Before discussing Jakobson's attitude towards phonological 'zero' it is necessary to examine Haas' treatment of the concept in phonology (Haas 1957). Haas' initial claim (1957: 33) is that

Zero in Linguistic Description stands for what is acoustically nothing.

This immediately appears to dictate that, for Haas, not only is "what is acoustically nothing" hermeneutically self-evident but also any 'overt' acoustic (semantically relevant) contrasts must be recorded as phonemic. For instance, German

(word initial) [ʔ] must be a phoneme. This turns out to be the case (ibid.: 36). Typically, Haas' more specific concerns are with zero in morphology, and with the concept of juncture. He attempts to impose strict constraints on the use of zero in linguistic analysis, particularly in phonology. In fact, there can be no place for it (ibid.: 36). Haas characterizes zero *qua* open juncture as a "specious zero element" (ibid.: 37) and states that:

There must be something amiss with a method of analysis that allows us to replace perceptible distinctions with imperceptible ones.

However, Haas at no time provides a clear statement of what is linguistically perceptible, and his attitude is essentially moralistic, as evidenced by such statements (1957: 36–37) as "The *abuse* has not been confined ...", "It is *ominously capable of extensive application*", and the already quoted "There *must be something amiss* ..." [My emphasis, WR].

Haas approaches the problem with two basic assumptions.

- i. A linguistic element must have a *form* of its own, identifiable as 'the same' in different environments (= static, non-parametric *class*)

and

- ii. it must have *distinctive value*, i.e., contrast with others in some of its environments (ibid.: 41).

Haas further states that "the point about zero is that it is identifiable *only by its distinctive value*". This, however, is merely an apriori assumption, and the alternative hypothesis, that some kind of phonic features are identifiable with the zero values postulated in phonology (requiring the kind of reversal of the usual analytical procedure mentioned by Hockett (1955: 156), is equally valid. This applies across all linguistic levels/units, including morphology. Haas' acceptance of so-called morphological zero is merely an epiphenomenon of the 'meaningfulness' of this unit, i.e., of the acceptance of the arbitrariness of the linguistic sign. Thus, the correctness of the claim is true *only* if we accept Haas' first premise, and that premise firstly is a highly empiric one, and secondly attaches priority to systems of 'overt' elements over the structure within which the systems operate, i.e., the paradigmatic has priority over, and *in all cases* dictates, the syntagmatic. This disallows completely the prior abstraction of structural templates and denies their priority. Robins (1957), albeit within a similarly highly empiric observationalist framework, regards structure as more immediately abstractable and as

allowing more general linguistic statements. Such a disagreement in fact argues only for the reduction of this dualism.

This is not to say that Haas' admonitions do not have value in many instances. For instance, the suggested constraint that (1957: 41) "zero itself can contrast *only* with an overt element, never with acoustic zero" is at least worthy of consideration, for it raises the serious issue of 'overtness', 'discernibility', etc. in the process of (linguistic) theopoesis. Moreover, his suggestions (1957: 42) that:

One may decide to regard acoustic naught as an empty segmental position (i.e., zero structural place)
and
zero, we might say, is not a 'segment' supplied; it is some determinate segment (or a class of such) *not* supplied. It is an *operation* performed on overt forms.

are not uninteresting except in the sense that zero is dealt with as a kind of 'voiding', as in Harris' terms (1951: Chap. 8.; cf. also Pāṇini's *lopa ve*). However, there is no solid theoretic reason to make 'zero' a *negative* product and 'non-zero' forms positive ones. This issue again is about dualism, particularly the syntagmatic vs. the paradigmatic axis, of which segment vs. place is merely a restatement. Yet the dualisms rampant throughout linguistics — and all other disciplines — are *all* reducible to the primal mind:body dualism. Within a linguistic program, a 'zero value' choice may be as 'positive' — i.e., as much a quantum — as the non-firing or inhibition of a neuron may be in a neural program.

Haas ultimately chooses to adopt the position that there are no good reasons for establishing *phonological* zero-elements, and that zero can have no phonemic features or allophones. Thus, the apparent contrast *stable*–*table* does not involve a zero phoneme, i.e. #[s] vs. #[Ø] opposition, and *stable*–*table* is only a putative 'minimal pair'. From this it must follow that, since *stable* differs from *table* in terms of the number of phonemes, the difference is a *structural* rather than a *systemic* one, and so the linguistic element word is not a singular notion but a set theoretic one. This notion, when generalized, makes all linguistic elements, and hence all linguistic levels, set theoretic! For Haas, as in linguistics in general, structure is an epiphenomenon of the linearization/distribution of the systemic elements phonemes, and the non-singulariness of linguistic elements/levels is not bothersome.

Working within the confines of phonemic tradition, it is not surprising that Haas arrives at the conclusion that he does. Within phonemics — Trubetzkoyan (Praguian), American, or Jonesian (1950) — the fact that zero 'has no allophones' by analytic *fiat* is sufficient criticism, or at least a sufficient foreclosure,

however theoretically strange and inconsistent this statement might be. However, it is not by proving the value of 'zero' in phonemics that its worth in linguistic theory is to be demonstrated, but rather in an entirely new conception of theory.

2.2 *Zero from Jakobson to TGP*

Jakobson, on the other hand, has no problem with postulating a zero phoneme (Jakobson & Lotz 1949: 431):

A ZERO PHONEME, which may be symbolized by \emptyset , or, in an analytical transcription, by #, is opposed to all other French phonemes by the absence both of distinctive features and of a contrast sound characteristic. On the other hand, the zero phoneme \emptyset is opposed to the absence of any phoneme whatsoever. In the initial prevocalic position this phoneme is known under the name 'h aspiré'; although under emphasis it can be implemented as an aspiration, usually it is a lack of sound, which acts in the given sequence as do the French consonants. The vocalic variant of the zero-phoneme, which appears in the other position, is called 'e caduc' and alternates between the presence and absence of a vowel.

Jakobson's zero phoneme then is not a juncture, but rather a 'glide type' phonemic element. Its very nature — in that it is, under specific style conditions, 'acoustic zero' in Haas' terms — would be more than suspect to the latter and be completely unacceptable. It is to be noted, however, that Jakobson's French zero phoneme is overtly realizable under other stylistic conditions where it is not acoustic zero. Thus, the stylistic problem and the question of generalization for 'language' is clearly at issue here. Clearly, also, the postulation of such a 'zero' makes more sense and is less problematic morphophonemically.

Jakobson (with Fant and Halle) (Jakobson et al. 1951: 39, 43ff.) postulates a 'zero' consonant phoneme for English, as per the tradition of the Greek and Roman grammarians:

The prevocalic or post vocalic aspiration /h/ is opposed to the even unaspirated overt or decay of a vowel. The former is a tense glide (*spiritus asper*), the latter, a lax glide (*spiritus lenis*), which properly speaking is a zero phoneme. This opposition (/h/ — /#/) occurs in English in initial prevocalic position: *hill* : *ill* - *pill* : *bill* ...

In adopting this approach, Jakobson brings the contrast down to the segmental level from the prosodic one at which the contrast had been handled from the period of the Alexandrian orthographers, a tradition endorsed by the Latin grammarians (Cf. Quintilian's "*h non littera est sed anima vocis*").

Within Jakobsonian DF theory another version of the zero-issue enters the picture in that the presence of a feature [+] may contrast with its so-called absence [-]. It is possible, thus, that a positive attitude towards zero could arise from a binarist approach to phonology, where minus values are very often automatically associated with 'absence' of something. There is a similar concept utilizable in the sphere of segmental as opposed to 'feature phonemes'. Thus Hockett (1955: 66) states

In the case of a two-term [linear] system [with zero] we can define one of the two as involving the *presence* of a particular structural unit, and the other as marked by the *absence* of that same sound.

It is clear, however, that we are dealing in this case with a notational issue, viz. notational economy, where a 'non-mark' may be both contrastive and have clear phonetic correlates different from those of the 'marked' element. As Hockett points out (1955: 69).

From a purely logical point of view it would be possible, in these systems ... to 'zero out' one of the contrasting terms ... But we must distinguish between possible notational devices and the phonologic pattern which is represented by them. In a linear system without zero there is no structural basis for deciding which terms are positive, which term negative; one term of the set is just as positive as another. This is all the meaning that need be read into our labels 'with zero' and 'without zero'.

Thus, apart from Chao's (1934) proposal, the concept of zero vowel is completely alien in linguistics. Allen (1956) and Kuipers (1960, 1969, 1970) never mention the possibility of a /Ø/ vowel phoneme — however much of this may be implied in the statement 'absence of *a* implies *ə*' (Allen 1965: 114). In Abaza and in Kabardian, all syllables begin with a [C], so a zero vowel is clearly not an impossible concept. That it is hardly ever considered is due to the requirement of a(n overt) vowel as nucleus of the unit syllable.

3. Phonetics and linguistics (phonology)

The issue of zero is obviously closely bound to the conceptualization of the two levels of scientific discipline — phonetics and linguistics (more specifically, phonology) — and of the relationship between them. On this, the general position is and continues to be what can be called the *phonetic realism principle* (PR(P)), this despite the fact that it should be equally obvious that phonetics is no less theoretic, or abstract, than phonology, and that the phonetics to which

linguistics is tied via the PRP is itself driven by the semantic basis of linguistics.

The one serious attempt at supplanting the PRP is to be found in Bloomfield's (1933; cf. also Fudge 1967, 1969) attempt at a full-blown distributional linguistics, brilliantly expounded in Harris (1951). One of the unnoticed conclusions of Bloomfield's approach is that it shows that all phonemes are distributionally unique. The failure to take note of this, and to extend it to an ordering of phonemes, by sub-classes and within sub-classes, leads to the teleological notion of natural class and natural rule, central in TGP, which in turn creates an inappropriate evaluation metric. Strong reliance on realist phoneticism was a distinguishing feature of post-Bloomfieldian phonology. The emphasis on relativity in defining phonological elements, characteristic of Sapir, Saussure, Firth, et al. (cf. Sapir 1925: 203) more or less disappears and the observationalist, empiricist confidence in phonetics as generally of immediate linguistic relevance continues through the present. It is still possible to say that, for all schools, the principle of invariance is the guiding one (Chomsky & Halle 1968: 168, 295–298). The position of post-Bloomfieldian linguistics is clearly stated in Hockett (1955: 156):

Segments which are in complementation are nevertheless phonetically distinct unless they resemble each other phonetically in such a way that the differences can be extracted from the segments themselves and assigned, instead, to the environments, in some phonetically realistic way.

This strong Realist attitude towards phonetics bars the possible identification of, for example, onset [k] and peak [a], but would certainly allow an identification of onset (or coda, or complex-peak satellite) [w] and peak [u]. It is clearly to Jakobson's credit (Jakobson et al. 1951) that, through the DF theory, he was able to create conditions under which [k] and [a] could be demonstrated to be more closely related than would appear in a phonemic analysis without DFs. Trubetzkoy's work on features also foreshadows this, but does not do it as clearly as the DFs do because he has different types of parametric relationships. Hockett was not unaware of the constraints imposed on linguistics by traditional phonetic technique (1951: 158):

It should hardly be necessary to add that failure to see phonetic similarity can be induced by bad phonetics, particularly by an uncritical reliance on traditional use of phonetic descriptive terms.

Phonology is presented in terms of the categories of a system of phonetics and the set-theoretic relations between phonological elements so defined are constrained by the criterial notions of phonetic similarity/dissimilarity (PS) they engender. Such criteria are purely static/aesthetic and very hard to 'define'. Yet, in phonemics, the criterion of phonetic similarity does in certain cases take

precedence over putatively formal criteria such as complementary distribution. Yet, while Bloch (1950: 108) frankly admits the adhocness of the PS criterion, Austin (1957: 538) poses the crucial question very bluntly:

Our science has reached a state where rigor is demanded in all its aspects. If rigorous rules of phonetic similarity cannot be deduced, let the criterion be dropped. Or are there rules for phonetic similarity that are universally applicable?

Austin's examination of some apparently drastic sound changes appears to indicate that the aperture vs. place-of-articulation matrix does, in general, represent a vaguely valid set of restrictions that account for the directionality of change along one axis or the other. While I too have found this, with extensive modification, to be the case, the issue of the reduction of these two axes — as well as the third axis generally used to characterize voice (glottal configuration) — never arises either in 'phonetics' or 'linguistics'.

Earlier work in TGP did not operate with an explicit PRP. Thus, Halle (1958 [1964]: 332):

As already noted, the rules that constitute the phonological component of a grammar relate a matrix consisting of abstract markers — the phonemic representation — to a matrix where each marker represents a particular aspect of vocal tract behavior. The latter matrix is our counterpart of the conventional phonetic transcription. In the phonemic representation the different features are allowed to assume only two values, plus or minus. In this representation, however, *no phonetic content is associated directly with the features which function here as abstract differential markers.* [My emphasis, WR]

Later, this position was abandoned in favor of the Naturalness Condition (henceforth NC) (Postal 1968: especially 53), which is essentially merely a restatement of the *phonetic realism principle*, as a condition on the non-arbitrariness of the relation between the (systematic [morphophonemic]) phonemic and (systematic) phonetic levels. The NC has essentially the following form:

The phonetic representation of an utterance consists of a string of *phonetic segments*, each of which is a complex of independent multivalued features describing some aspect of the *ideal* behavior of the speech apparatus. Let us assume somewhat oversimply that the phonological representation of an utterance consists exhaustively of symbols of some kind. Then the task of the phonological rules is to map phonological strings onto phonetic ones. [My emphasis, WR]

The relation between phonological and phonetic structures is a *natural* one such that the form of representation for one is the same as for the other. This being

the case, the mapping process can to an unspecified degree be given universally (Postal 1968: 54, 67, fn. 15). Postal (1968: 56), repeating an essentially Jakobsonian position, imposes the following analogy on the situation of language learning:

When a child has determined the phonetic representation of a form, he has thereby determined a great part of its phonological structure.

Thus, the relation between phonological and phonetic structure (cf. Chomsky & Halle 1968: 294) is not an arbitrary one. A similar notion of the relation between the two levels is to be found in Firthian phonology, expressed in the statement that phonological structures "have the implication of utterance". Thus, the crucial principle is that phonological forms, rules, and constraints do not involve "randomly related phonetic mappings", and this is assumed to be universally the case (Postal 1968: 47, 58):

... it is immediately evident that in some strong, though far from completely clear sense, a natural relation between phonetic and systematic structure does obtain in all known languages ...

Postal opts for traditional segmentation and binary DF componentialization of segments (although no limit on the number of DFs is given). The degree of deviation from invariance in the one-to-one relationship between phonetic and phonological representations — and from biuniqueness and linearity — is a function of the number and kind of language-specific P-rules (all of which in turn constitute the universal inventory of P-rules). That is, it is a function of the requirements of establishing for the first a set of single, basic, static, invariant underlying representations (1968: 63), such that a one-segment to many relation and vice versa may hold between phonemic and phonetic representations, and that components and/or componential values at the systematic phonemic level may be different from those at the phonetic to the extent required for the generalized phonemic structure of morphemes. The derivation is 'correlated' by the P-rules in the mapping process. In addition to traditional segments and DFs, further information affecting the derivation and thus the form of the mapping process is required in the shape of grammatical information (boundary, categorical, etc.), which clearly involves another difficult and vague area in linguistics, viz. the relation between speech and grammar.

As with the PRP, there are several immediate and crucially important problems with the NC as developed by Postal (and also by Chomsky & Halle 1968). The one that has attracted the greatest amount of attraction (cf. Sommerstein 1977: 108–109) is the fact that the NC imposes "no limit to the complexity of the phonological rules a grammar may contain if the facts call for them". Put in a more common form, the NC imposes no limits on 'abstractness', specifically

derivational abstractness. The reason for this is obvious. Postal adopts the premise that language morpho(phono)logy in fact contains only two types of relationships — allomorphy and suppletion. Greater value is given to a grammar which derives the larger number of surface forms (allomorphs) by P-rule than to a grammar which contains a large number of suppletive form substitutions.

It is also possible to level another criticism at the NC which, if justified, has consequences of far greater import for linguistic theory than the above. The 'substantiveness' or 'intrinsicness' of the phonological elements postulated in Postal's version of the NC, is, as has already been stated, highly traditional. They are given apriori in the theory and established largely if not totally on the basis of contrasts noted in the languages of the world. In general, the segments are alphabetic and, since these allow structural statements in terms of C and V to be made quite conveniently, clearly and simply, very little consideration has been given to the possibility and feasibility of constructing a model in which the alphabetic segment plays no uniquely defined formal role, and in which C and V are unnecessary for the statements of structural generalizations. (On this, see further below.)

In the development of the NC and the statement of its primes, Postal, like all other generative phonologists, tacitly follows the strong exclusionary line on the issue of zero. Thus there is no zero morphophoneme, in the sense, for example, that *ill* is a two segment form and not a three or more segment form. Morphemes (formatives) contain, in general, no more and no fewer segments than are required by the Alternation Condition. No segments are postulated for the sake of *structural* generalization. There is only an option for the immediate *realist* economy over the eventual economy that would derive from postulating a zero element in phonetics and phonology.

Thus, the NC is merely a *Strong Phoneticity Requirement* [SPR] essentially no different from the PRP. The SPR dictates that the elements of the *technique* of phonetic description is essentially correct, i.e., psychologically valid (Postal 1968:passim). The problem of a formal theoretic condition linking a 'technique' to a 'theory' remains as opaque as always, however. Every deviation from the SPR is to be rigorously motivated by evidence from observed phonetic/linguistic processes. In this way, the NC via the SPR involves an in-built self-justification based on a highly empiric, procedural, descriptive approach to linguistics. It putatively starts with a technique of observation apriorily determined by the 'linguistic theory', which is then, as circuitously as in all other approaches to linguistics, elevated to the status of a theoretical condition, i.e., a condition on grammars insofar as it dictates the type of formal representations usable in a grammar.

4. Some proposals regarding zero in linguistic theory

4.1 *Linguistic forms*

In phonemics, canonical forms (CF), or shapes, use the brackets (), but the notions of () and CF as formal elements, e.g., PLACE and FORM, are non-existent. In TGP, at all levels, there is similar unclarity about the formal status of CF and there is no such thing as a 'zero' systematic phoneme, i.e., *ill* is a *two matrix* form, *hill* a *three matrix* one, etc. It is a general condition on elements enclosed with () that none of them can have a zero value. If this were the case, then () would obviously be unnecessary. In this sense () states that some *non-zero* element(s) are *optional*.

The (un)bounded expression notation in TGP, e.g., C_0 or C^n_0 , at first glance appears to be the nearest thing to a linguistic zero element. It is, however, merely a notational variant for the parenthesis notation, which effects the conflation of rules (and, see further below, of linguistic level terms as sets). In a rule like $A \rightarrow B / _ (C) D$, neither ACD (nor non-systematic BCD) is a singular form, and while A, B, and D are non-zero, (C) is not zero but optional. Rules such as:

$$X \rightarrow \emptyset /$$

$$\emptyset \rightarrow X /$$

(which may be context free or context bound rules)

are *interlevel* rules, and essentially this position is not different from the structuralist morphophonemic convention. Thus, both the unbounded expression, the parenthesis notation, and the zeroization rules are merely means of stating cross-level correlations between set representations — some higher level non-zero element being voided at a lower level. In TGP, \emptyset has no elemental function, i.e., it is not a form.

All phrase structure grammar (PSG) conventions are 'canonical', and in TG have the added property of recursion — i.e., every unit may consist of its components and itself (theoretically an infinite number of times). Traditionally, the CF notation has most commonly been used to express syllable structure without the possibility of recursion. This constraint reflects the difference between 'meaningless' (phonological) units which are of finite size and 'meaningful' elements whose size is not theoretically bounded.

4.2 *Syllable structure*

Abercrombie (1967: 73–75) discusses the form of the English syllable in the following manner:

PLACES	1	2	3
<i>go</i>	C	V	O
<i>oat</i>	O	V	C

O, or zero,

indicates that place 1 or place 3 is *empty*, with no element of structure there. Thus the four syllables just quoted [sc. *oh*, *go*, *oat*, *goat*: WR] exemplify the following four *different patterns* [OVC, CVO, OVC, CVC: WR]. [My emphasis, WR]

... These ... patterns of syllable structure can be expressed — or compressed — in a *single generalized formula* $C\ V\ (C)$, where the brackets around the *C* indicate the optional presence of an element at that place — (C) is equivalent to 'either C or O'. [My emphasis, WR]

A generalized formula for English syllable structure (ignoring the dubious *triumphst*), constructed in the same way as the preceding ones, would be $(C)(C)(C)V(C)(C)(C)(C)$. However, this is neither very elegant nor compact. There is a simpler way of generalizing syllable structure formulas for languages which have clusters. The formula for English, for example, would be $C_{0-3}VC_{0-4}$. The subscript figures here indicate the possibilities, in terms of a number of elements for that place of syllable structure.

Note that the two statements — "O, ... indicates that [the] place ... is empty, with no element of structure there" and "where the brackets around the C indicate the optional presence of an element at that place — (C) is equivalent to 'either C or O'" — do not in fact make quite the same claim. The first says that there is no structural element/place, while the latter implies that there *is* a place! This notational use of O conflates the notions of 'zero' and 'optionality'. The only formal claim made here is that there is no formal zero element, and that a term such as a syllable is linguistically a set theoretic notion — i.e., it consists of a set of optional patterns. Note further that there is no discussion as to whether such patterns (set members) are ordered.

It is clearly the case that the unrestricted use of zero within the linguistic approaches available yields no principled manner in which smaller and larger unit tokens for what are regarded as meaning-bearing or grammatical units (morphemes, words, etc.) be distinguished. For any language, the smallest and the largest will be coterminous. There are the following undesirable consequences of this situation:

1. While there may be observational limits to unit size in any particular language, it appears undesirable to limit morpheme size on the basis of this.

2. If a uniform size corresponding to the largest observed meaning-bearing unit size in a language is established, then, at least according to the Chomskyan position, this would be, in principle, infinitely expandable. In the interest of generality, there is stronger motivation to establish the smallest size possible (for any grammatical class) both universally and in any particular language and, thus, to attempt to establish universal principles of expansion. The clearest candidate is finite or non-finite recursion.

Just as clearly, it is not required of a theory that it establish apriori operational procedures for defining/discovering its units — its primitives. For some, it may be sufficient to say *what* it is — e.g., the universal syllable structure is /CV/. However, this generally amounts to nothing more than a broad experiential *qua* typological extrapolation, which, if it is intuitively satisfactory, becomes a perpetuated linguistic canon.

4.3 *Zero and the C:V dichotomy*

As is evident, the postulation of zero-element and the C-V dichotomy are also very closely related (Roberts 1972a). Furthermore, the use of zero involves crucially the syntagmatic-paradigmatic duality and ultimately the whole fabric of phonetic classification. Although Jakobson does not examine these issues, it is not without significance that the use of zero and of overlapping descriptions (DFs) for both consonants and vowels co-occur in his phonological approach. The conflicts between theoretical premises, analytical procedures, and the demands of 'phonetic realism' are at all times involved in the issue, and confusion of levels of scientific discourse all too easily nurtured. The characterization of the use of zero as a *deus ex machina* (cf. Mulder 1968: 118) is all too often based on the demands of *credo* and nothing more. It should be kept in mind, however, as Hockett (1955: 158) pertinently admonishes us, that existent phonetic systems have less than absolute criterial value. It has, to some degree or other, been accepted as axiomatic that the technique of phonetics is never applicable fully independently of linguistic functional considerations, conscious or unconscious. Indeed, "uncritical reliance on traditional use of phonetic descriptive terms" (Hockett loc.cit.) clearly leads to many failures in making or constructing linguistically relevant statements — in the areas, for example, of 'phonetic-phonological similarity', classification, and segmentation (Roberts 1972a, 1972b, 1996).

Despite their putative primary placement in the set of DFs, consonantality and vocalicity are, ironically, through the use of zero, the easiest and most immediate features to be shown to be predictable (i.e., redundant) as functions of

ORDER. If the distinction is a primitive formal one, then this should not be the case. This seems always to be a frightening possibility in linguistics, and this fear is well-evidenced in the occasional discussions that arise in the literature on the possibility of vowelless languages (Allen 1956; Kuipers 1960, 1969, 1970; Halle 1970). Counterarguments are always aprioristic (e.g., the phonetic realism principle), vaguely typological, and intuitive. For it must follow that if there are vowelless languages — and all linguistic theories, by virtue of being based on the principle of complementary distribution, allow the analysis of all languages to this extent — then what ‘remains’ are not consonantal languages but rather *language* in which the C:V dichotomy has no relevance!

From the above, it is obvious that there are no clear answers to questions such as: Why is there no phonological, ‘allophonic’, or phonetic zero element? Apart from the problems imported apriorily, there are logical procedural limitations and foreclosures which have prevented an intelligent approach to the issue and which in fact contradict the fundamental premise of linguistics — that it is about making or discovering generalities about speech-language. This can and has been variously expressed as a search for recurrent partials, for economy, and for making redundancy free statements (however this notion may be conceived).

When one considers a progression of pairs such as *Ide: ride; ride: tried, tried: stride*, then allowing the first and the last to be comparable as minimal pairs yields [ØØØaid] — and, by extension, [ØØraid], [Øtraid]. Possible arguments against this are:

- a. The physical (‘phonetic’) values of zero and zero in sequence may be judged impossible to establish, or to perceive. This is a typical apriori assumption and rejection of reasonable hypotheses extensively exemplified in the discussion above.
- b. Ø becomes a phoneme that has a unique distribution.
- c. Such a procedure would lead to the apriorily untenable position that the size of every linguistic element is equivalent to the largest exemplar (intra- or inter-linguistically). This seems inconceivable, and ‘redundant’.

Argument a. is based on typical apriori assumptions and on rejection of reasonable hypotheses as extensively exemplified in the discussion above. There is no clear reason for rejecting this as a physical (physiological, acoustic, perceptual) hypothesis. Argument b. is not as strong as it may appear, since, as was pointed out above, Bloomfield’s distributional definition of phonemes demonstrates that each phoneme has a unique distribution (a property not difficult to understand

from the point of view of sound changes). Argument c. is equally weak because it is based on premature foreclosure. The weakness is based on the failure to recognize a general pattern involving zero — that is, on a failure to concentrate on the extrapolation of recurrent patterns. This in turn is due to an apriori assumption of phonetic classification and representation, namely that, between so-called pauses *qua* suspension of speech, phonetic symbols represent non-zero vocal-tract (or impressionistic) configurations that are all in immediate contiguity. Phonetics does not allow a difference like the following: [pa] vs. [p.a], vs. [p..a], etc., where [.] indicates non-contiguity, which can then be interpreted as 'zero', and related to durational, intensity, and other speech properties.

It was stated above that, if one assumes that linguistic theoretic units, e.g., the SYLLABLE, are formally regarded as singular rather than as set-theoretic notions as is standard in linguistics, then every linguistic form, and synonymously, every linguistic level, is singular. There are no optionalities of constituency. If, for the sake of initial discussion, we assume that the SYLLABLE is theoretically (hence universally) a three-place constituency (in general phonetic terms consonant + liquid + vowel), and each place has potentially zero-value, then the distinction between consonantality, liquidness, and vocalicity becomes predictable as a function of linear order. This has clear advantages for extending the concept of (phonetic) similarity (well beyond Jakobson's basis for establishing the [acoustic] DFS) and increasing the generality (while automatically decreasing the redundancy) of prevailing linguistic theories. Admittedly, the common paradigm-phonetics based notions of syllable, cluster, etc. will have to change.

Furthermore, the assumption of a triadic structure for SYLLABLE suggests an extension of triadicity within the theory. Note that the structure of the syllable is representable as, e.g., xxx, or •••, which means that the structure is recursive to the quantum 3. Each quantum is a 'segment' at the SYLLABLE level. Let us hypothesize further that each constituent ('segment') of the constituency SYLLABLE is also triadic in structure. Thus, syllabic • consists of •••, i.e., each syllable segment also consists recursively of an equivalent triad. Let us assume further that the speech correlates of these three are, in order, LABIAL (cavity front of the exterior facets of the teeth), LINGUAL (cavity of tongue activity between the interior facets of the teeth and the palatal-velar junction), and LARYNGEAL (cavity from velum to the larynx). These correspond to the standard PLACE-OF-ARTICULATION (POA) designation (applied to 'consonants', 'liquids', and 'vowels'). Finally, assume that this second recursion has a third equivalent triadic recursion for each 'place' where these constituents correspond in order to the standard APERTURE-OF-ARTICULATION (AOA) designation (applied to 'consonants', 'liquids', and 'vowels') to, say, HIGH, MID, and LOW. Note that so-called phonetic tokens of the unit/levels syllable/

segment, and POA/AOA features are not necessarily accounted for by a single triad. A 'syllable'/'segment', POA/AOA feature may be $\bullet\bullet\bullet$, or $\bullet\bullet\bullet\ \bullet\bullet\bullet\ \bullet\bullet\bullet$, etc. within the limits of the triadic recursion. This then accounts for the phonic variation at these levels. In this sense, the theory provides a predictably *finite* linguistic amount of phonic-linguistic variation, within and between languages. Thus, by showing *typology* to be only *statistically relevant at a relatively highly empirical and qualitatively varied level of discourse*, and to be *linguistically irrelevant*, the theory is truly *universal*.

Assuming that this is the last recursion, the consequences should now be obvious: there are three recursions — the first is predictably segmental (a syllable is three places) and the general phonetic class of the places are predictable; the second is predictably place-of-articulation (a segment is three place-of-articulation places) and the general phonetic class of the places is predictable; the third is predictably aperture-of-articulation (a place-of-articulation place is three aperture-of-articulation places) and the general phonetic class of the places are again predictable. Thus, with a recursive model of this sort, all structures (levels) are singular, all structures (levels) are linearly ordered, all standard phonetic classifications (with appropriate respecifications) are predictable (syllabicity, consonantality and vocalicity, segmentality, place-of-articulation, aperture of articulation). Finally, the syntagmatic:paradigmatic (structure:system) dichotomy is reduced, binariness is a local (i.e., relative) function of recursion place value, and the arbitrariness of the linguistic sign hypothesis becomes linguistically irrelevant (cf. Toussaint 1983). Thus, we have a(n autonomous) linguistic theory. This fits nicely and simply with the notion that speaking is, after all, nothing more than variation in the process of opening and closing the mouth. The extension of this form of recursiveness 'upwards' in triads from the three-segment (syllable) should be clear, yielding a quantal theory of the unit sentence as ([3 to the power of three] to the power of three). This requires that serious changes of conception must take place, for instance, that 'morphemes' do consist of 'syllables'. It also follows that linear order is itself predictable from recursion.

5. Concluding remarks

The introduction into linguistic theory of an element or concept which, after long and serious consideration by the members of the discipline, is less than generally acceptable is a serious matter. The arguments and counter-arguments must be carefully examined and the adoption of the element must then be shown to contribute significantly in the area of the generality of statements possible in its

presence, impossible in its absence. Furthermore, any demonstrated ensuing generality must itself be acceptable. This is a common matter of theoropoesis — of science in its most general sense. Of course, there is no guarantee of acceptability. This is the price of the effort — that everything is subject in the end to a matter of credibility. With the proposal outlined here for linguistic zero as a place where there is minimal value, or energy, all the problems that have plagued it within linguistics can be seen to disappear. The advantages of generality and maximum predictability are clearly present in this proposal. The question now is whether there is sufficient *tolerance* to accept it.

Author's address

E. Wyn Roberts
 Department of Linguistics
 Simon Fraser University
 Burnaby, British Columbia, CANADA V5A 1S6
 Wyn_Roberts@sfu.ca

References

- Abercrombie, David. 1967. *Elements of General Phonetics*. Chicago: Aldine.
- Allen, W. Sidney. 1955. "Zero and Pāṇini". *Indian Linguistics* 16.106–113.
- . 1956. "System and Structure in the Abaza Verbal Complex". *TPS*. 127–176.
- . 1965. "On One-vowel Systems". *Lingua* 13.111–124.
- . 1962. *Sandhi: The theoretical, phonetical and historical basis of word-junction in Sanskrit*. The Hague: Mouton.
- Austin, William M. 1957. "Criteria for Phonetic Similarity". *Language* 33.538–544.
- Bloch, Bernard. 1950. "Studies in Colloquial Japanese IV". *Readings in Linguistics I* 1957 ed. by M. Joos, 329–348. Chicago: University of Chicago Press.
- & George L. Trager. 1942. *Outline of Linguistic Analysis*. Linguistic Society of America. Baltimore: Waverly Press.
- Bloomfield, Leonard. 1926. "A Set of Postulates for the Science of Language". *Readings in Linguistics I* 1957 ed. by M. Joos, 26–31. Chicago: University of Chicago Press.
- . 1933. *Language*. London: George Allen & Unwin.
- Chao, Yuen-Ren. 1934. "The Non-uniqueness of Phonemic Solutions of Phonetic Systems". *Readings in Linguistics I* 1957 ed. by M. Joos, 38–54. Chicago: University of Chicago Press.
- Chomsky, Noam & Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper & Row.
- Conant, James B. 1957. *Science and Common Sense*. New Haven: Yale University Press.
- Fudge, Eric. 1967. "The Nature of Phonological Primes". *Journal of Linguistics* 3:1.1–36.

- . 1969. "Mutation Rules and Ordering in Phonology". *Journal of Linguistics* 5:1.23–38.
- Gleason, H. Allan, Jr. 1961. *An Introduction to Descriptive Linguistics*. 2nd ed. New York: Holt, Rinehart & Winston.
- Haas, Werner. 1954. "On Defining Linguistic Units" *TPS*. 54–84.
- . 1957. "Zero in Linguistic Description". *Studies in Linguistic Analysis* ed. by J. R. Firth, 33–53. Oxford: Blackwell.
- Halle, Morris. 1958. "On the Bases of Phonology". *The Structure of Language: Readings in the philosophy of language* 1964 ed. by Jerry A. Fodor & Jerrold J. Katz, 324–333. Englewood Cliffs, N.J.: Prentice-Hall.
- . 1970. "Is Kabardian a Vowel-less Language?". *FL* 6:1.95–103.
- Harris, Zellig S. 1951. *Structural Linguistics*. Chicago: University of Chicago Press.
- Hartman, Lawton M. 1944. "The Segmental Phonemes of the Peiping Dialect". *Readings in Linguistics I* 1957 ed. by M. Joos, 116–123. Chicago: University of Chicago Press.
- Hill, Archibald A. 1958. *Introduction to Linguistic Structures*. New York: Harcourt Brace.
- . 1966. *The Promises and Limitations of the Newest Type of Grammatical Analysis*. Cincinnati: University of Cincinnati.
- Hockett, Charles F. 1947. "Peiping Phonology". *Readings in Linguistics I* 1957 ed. by Martin Joos, 217–228. Chicago: University of Chicago Press.
- . 1955. *A Manual of Phonology*. *Indiana University Publications in Anthropology and Linguistics. Memoir II. Pt. 1 of IJAL* 21. Baltimore: Waverly Press.
- Jakobson, Roman, Gunnar Fant & Morris Halle. 1951. *Preliminaries to Speech Analysis*. Cambridge, Mass.: M.I.T. Press.
- & John Lotz. 1949 [1964]. "Notes on the French Phonemic Pattern". *Roman Jakobson. Selected Writings I* 1962 ed. by R. Jakobson, 426–434. The Hague: Mouton.
- Jones, Daniel. 1950. *The Phoneme: Its nature and use*. Cambridge: Heffer.
- Joos, Martin ed. 1957. *Readings in Linguistics I*. Chicago: University of Chicago Press.
- Kuipers, Aert. 1960. *Phoneme and Morpheme in Kabardian*. The Hague: Mouton.
- . 1969. "Unique Types and Phonological Universals". *Pratidanam. Volume in Honour of F. B. J. Kuiper*. The Hague: Mouton.
- . 1970. "Linguistic Truths and True Trivialities: M. Halle on Kabardian". mimeograph (ed.)
- Moulton, William G. 1947. "Juncture in Modern Standard German". *Readings in Linguistics I* 1957 ed. by Martin Joos, 208–215. Chicago: University of Chicago Press.
- Mulder, Jan W.F. 1968. *Sets and Relations in Phonology*. Oxford: Clarendon Press.
- Nida, Eugene. 1947. *Morphology*. Ann Arbor: University of Michigan Press.
- Postal, Paul M. 1968. *Aspects of Phonological Theory*. New York: Harper & Row.
- Roberts, E. Wyn. 1972a. "Consonant and Vowel: A re-examination". *Lingua* 30.141–212.
- . 1972b. "A Critical Survey of Firthian Phonology". *Glossa* 6.3–73.

- . 1996. "Recursion, Linear Order, and the Classification of Linguistic Units". *Typology: Prototypes, item orderings and universals. Proceedings of LP '96* ed. by Bohumil Palek, 11–124. Prague: Charles University Press.
- Robins, R. H. 1957. "Aspects of Prosodic Analysis". *Proceedings of the University of Durham Philosophical Society* 1.1–12.
- . 1964. *General Linguistics. An introductory survey*. London: Longman.
- Sapir, Edward. 1925. "Sound Patterns in Language". *Readings in Linguistics I* 1957 ed. by Martin Joos, 19–25. Chicago: University of Chicago Press.
- Saussure, Ferdinand de. 1915. *Cours de linguistique générale*. 1st ed. 1955 (5th ed.). Paris: Payot.
- Sommerstein, Alan H. 1977. *Modern Phonology*. Baltimore: University Park Press.
- Swadesh, Morris. 1934. "The Phonemic Principle". *Readings in Linguistics I* 1957 ed. by M. Joos, 32–37. Chicago: University of Chicago Press.
- Toussaint, Marcel. 1983. *Contre l'arbitraire du signe*. Paris: Didier.
- Trubetzkoy, Nikolai S. 1939. *Grundzüge der Phonologie*. Göttingen: Vandenhoeck & Ruprecht. (= *Principles of Phonology* 1969 transl. by Christiane A. M. Baltaxe). Berkeley & Los Angeles: University of California Press.
- Twaddell, W. Freeman. 1935. "On Defining the Phoneme". *Readings in Linguistics I* 1957 ed. by Martin Joos, 55–79. Chicago: University of Chicago Press.

CHAPTER 25

'God's Truth' and Structuralism

A new look at an old controversy

Gary D. Prideaux
University of Alberta

1. Introduction

The transition from a historical-comparative perspective on language study to a structuralist approach brings with it, virtually by necessity, a commitment not only to the investigation of language structure but also a concomitant commitment to an analysis of the human cognitive mechanisms involved in language use. The specific thesis developed here is that the structuralist constructs of constituent structure and markedness are not merely convenient notations but in addition have firm empirical support in language use.

Each period in our intellectual history somehow defines itself, at least implicitly, in terms of a particular metaphor. The metaphor associated with the historical-comparative period was, of course, the Darwinian notion of the 'family tree' and of genetic relations, a useful metaphor which persists to this day, finding itself somehow re-defined and revised in ways congruent with the advances in the biological sciences. The structuralist period struggled to define for itself a clear philosophical perspective, but at least a rough picture seems clear: it was, at least in North America, basically anti-mentalist, drawing upon the behaviorism of the day (though it is clear that such a characterization is perhaps more of a caricature than a clear picture, since it could be argued that some of the most influential linguists of the period (e.g., Sapir and his students) did not entirely subscribe to Bloomfield's (1933) strong behaviorist position). The structuralist metaphor seems to have been that language is mindless, independent of its speakers and their minds.

One useful way to characterize some of the central controversies of the structuralist period, especially in the North American context, is to focus on the

question of whether the linguists' analyses were 'convenient fictions' (e.g., Twadell 1935) or were indeed somehow representative of what speakers seemed to know. Perhaps this controversy was best articulated by Householder's 'God's truth' versus 'hocus-pocus' dichotomy (see, e.g., Joos 1958: 80). Hockett (1948), for one, appeared to support some version of the 'God's truth' position. In his influential 1948 paper, Hockett criticized the view that linguistic analysis is just a game, and in particular when he rejected Preston's (1948: 131) assertion that "The structure of a given language or language corpus does not exist until it is stated". In particular, in rejecting the *hocus-pocus* view, Hockett (1948: 269) asserted:

The task of the structural linguist, as a scientist, is, as Preston implies, essentially one of classification. The purpose, however, is not simply to account for all the utterances which comprise his corpus at a given time... Rather, the analysis of the linguistic *scientist* is to be of such a nature that the linguist can account also for utterances which are *not* in his corpus... That is, ... he must be able to predict what other utterances the speakers of the language might produce... (original emphasis).

Moreover, Hockett goes on to compare and even equate the linguist's task with that of the child acquiring language. This is a point, of course, which he extends and elaborates in his later book (Hockett 1957). Similarly, Sapir (1925) argued for the 'psychological reality' of the phoneme when he demonstrated that phonology as a system could not be understood in simple 'mechanical' terms. In contrast to the 'God's truth' position, other influential linguists of the period adopted the behaviorist position that linguistic descriptions are no more than 'convenient fictions'. Twadell (1935), for example, claimed that the phoneme is an "abstractional, fictitious unit".

In the domain of syntax, however, the structuralists were rather less vocal. Harris' (1946) rigorous procedures for defining substitution classes, along with Wells' (1947) proposals for immediate constituent analysis, became the source of later syntactic constituent structure descriptions, albeit in very different guises. Yet these scholars were silent as to the psychological content of their proposals, and it is even possible that questions of psychological reality would have made no scientific sense to them. One who was not silent was Martin Joos (1950), who adamantly maintained that linguistics was a descriptive and not an explanatory science. When contrasting the mentalistic views of Sapir with the behaviorist position of Bloomfield, Joos (1958: v) gently scolds Sapir's search for a mentalistic foundation for linguistics by asking "...why borrow trouble by explaining the invisible?"

More recently, the cognitive (and/or computational) metaphor seems to be

the one which captures the spirit of much current research in linguistics. For many linguists today, linguistic structure only makes sense in terms of a cognitive orientation, with a strong related computational spin (and why not call this a 'God's truth' approach as well?). A vast body of literature has emerged over the past decades as a result of trying to make empirical (psychological or cognitive) sense of linguistic theories, with varying degrees of success (for an excellent review, see Gernsbacher 1994). However, many structuralist contributions also came to play important roles in the later development of generative grammar. Harris' (1951) transformations are a classic and well-documented example.

Two other structuralist contributions which have taken on a new life within the modern frameworks are constituent structure analysis and markedness theory. While these originally served primarily as descriptive devices, they have been incorporated into, and have come to play important roles within, the development of generative grammar in its many guises. Moreover, the empirical status of these notions came to be seriously investigated once issues of psychological reality, mental processing, and mental representations became important concerns.

In the present paper, I wish to offer a modest contribution in this particular area by providing some examples from my own work in psycholinguistics and discourse analysis which reveal the necessity of constituent structure and markedness to explain certain phenomena found in language production.¹ Far from being entirely arbitrary, these theoretical constructs have significant empirical content and provide us with tools for further insight into how we process and represent language.

2. Constituent structure and closure

When psycholinguists began to turn their attention to language processing in the late 1960s, the role of constituent structure naturally played a central role. Once the (in)famous "derivational theory of complexity" was abandoned, several new proposals dealing with linguistic structure were formulated to deal with language processing. One of the most influential, and one which received considerable attention and reformulation over the next decade or so, was the principle of closure (see, e.g., Kimball 1973; Frazier & Fodor 1978; Frazier 1985; Prideaux & Baker 1986). The principle of closure can be formulated roughly as follows:

CLOSURE: A constituent (e.g., clause) which is internally interrupted by another instance of the same constituent (e.g., another clause) will require more processing resources (i.e., will be more difficult to process) than that same constituent which is not internally interrupted.

The first problem in dealing with an empirical assessment of closure is that of determining what it is that is closed. Many psycholinguistic arguments have been offered in support of the psychological reality of constituent structure, at least to the extent that it correctly predicts syntactic chunking. Similarly, closure can only be formulated if we have a clear and consistent notion of constituent structure which involves both hierarchic organization and category membership.

Closure is a generalized Gestalt principle in terms of which we impose coherent structure on our linguistic experience, a principle by which we package information. Because of working memory limitations, for example, we have only limited resources to allocate to a task and closure is a means for assembling one piece of information, perhaps at the same time clearing our active (syntactic) buffer in order to deal with new incoming material. Although several versions of closure have been offered, they all predict that an uninterrupted structural unit (clause, phrase, etc.) should be easier to process than one which is interrupted, simply because suspension of processing on one unit in order to attend to another increases the demands on working memory.

Most of the evidence offered in support of closure has focused on its role in language comprehension, with considerably less attention to its role in production. It is precisely to the issue of production that the first study is devoted.

English sentences containing relative clauses (RCs) provide a convenient testing ground for closure, since English sanctions RCs within NPs playing virtually any syntactic role. Closure predicts that sentences containing internal (interrupting) RCs should be in general harder to process than those containing non-interrupting (final) RCs. Clearly this prediction is one of tendencies and not absolute rules. While a sentence like (1a) is as grammatical as (1b), closure predicts that (1a) should be more difficult to process than (1b).

- (1) a. The guy who stole the bike frightened the electrician.
- b. The electrician frightened the guy who stole the bike.

An operational definition of processing ease can be formulated in terms of the relative frequency of occurrence within a written or oral text. According to closure, the non-interrupted structures (which should be easier to process) should be more frequent than the relatively harder (interrupted) structures. Thus, sentences such as (1b), with a final RC, should be more frequent than those like

(1a), which contain an internal RC. For the purposes of our study, we construct the closure hypothesis as follows:

CLOSURE HYPOTHESIS: proportionally more final than internal RCs should be found in (written and oral) discourse.

Our research group carried out two experiments to test this hypothesis, one dealing with written and one with oral narrative descriptions. In both experiments, participants watched a silent video tape and then provided a narrative description, either written (Experiment 1) or oral (Experiment 2), of what they had seen. Twenty-four participants volunteered for Experiment 1 and 16 for Experiment 2. Participants in both experiments were told to assume that the addressee (reader or hearer) had not seen the clip, so their task was to provide an extensive enough description that the addressee would have a clear understanding of what had taken place. Participants in Experiment 1 watched the clip in groups and then wrote their descriptions, while those in Experiment 2 watched the clip singly (in the absence of the experimenter) and then narrated their descriptions to the experimenter who taped their narratives for later transcription.

Each oral narrative was transcribed in conventional orthography, complete with pauses, false starts, hesitations, etc. As expected, the oral data differed greatly from the written in a variety of ways. Within each narrative, all RCs were coded as to type (restrictive or non-restrictive), location (clause-internal or final), and a variety of other factors which are of no immediate concern to us here. The pooled written data yielded a total of 90 restrictive RCs and 34 non-restrictives. The results from Experiment 1 are found in Table 1.

Type	Final	Internal
Restrictive	73	17
Non-Restrictive	24	10

Table 1. Written relative clause distributions

From Table 1 it can be seen that there are far more final than internal restrictive RCs (73 versus 17). However, a further factor must be considered, namely that English provides more opportunities for internal than final RCs, since the language allows more opportunities for host NPs within a sentence than at the end. To take account of this, the number of potential internal and final RC host NPs was counted for every clause containing a RC. Of the 251 potential host NPs in these clauses, 159 (63.3%) were internal, while 91 (36.7%) were final. These proportions were taken into account when testing for the strength of the

location effect. The expected internal and final values for the RCs under the null hypothesis are therefore the total number of RCs properly weighted for internal or final positions, yielding $E_i = 57.0$ and $E_f = 33.0$. Table 2 represents the expected and observed frequencies for the restrictive RCs.

Position	Expected	Observed
internal	57.0	17
final	33.0	73

Table 2. Observed and expected restrictive oral RC distributions

A χ^2 -test revealed that the position effect was highly significant ($\chi^2(1) = 76.55$, $p < .001$), indicating a strong tendency for restrictive RCs to be attached to final NPs, thereby strongly supporting closure. A similar analysis for the non-restrictive RCs also yielded a significant position effect ($\chi^2(1) = 3.97$, $p < .05$).

Exactly the same stimulus material was used for the oral experiment, in which 16 native speakers participated as volunteers, and precisely the same kinds of information were tabulated in this experiment as in the first. In the oral data, a total of 114 restrictive RCs was found, but only 8 non-restrictives. The results are found in Table 3.

Type	Final	Internal
Restrictive	61	53
Non-Restrictive	6	2

Table 3. Oral relative clause distributions

In testing for closure in the oral data, the potential position effect was again taken into account. Of the total 346 potential host NPs within the clauses containing restrictive RCs, 256 (74.0%) were internal, and 90 (26.0%) were final. The expected internal and final values for the RCs under the null hypothesis of no differences are therefore $E_i = 84.4$ and $E_f = 29.6$. Table 4 represents the observed and expected frequencies, upon which a χ^2 -test can be performed.

Position	Expected	Observed
internal	84.4	53
final	29.6	61

Table 4. Observed and expected restrictive written RC distributions

A χ^2 -test ($\chi^2(1) = 44.99$, $p < .0001$) revealed a significant tendency for restrictive RCs to be found in final position, again supporting closure, and while the number of non-restrictive RCs in this sample is really too small for a clear interpretation, their proportions also point toward the support of closure. These results not only provide strong support for the operation of closure in language production but, more to the present point, they also support the importance of a cognitive interpretation of constituent structure for understanding how speakers actually structure their language in production.

3. Constituent structure and markedness

We now turn to markedness, a second area in which structuralist analysis has made important contributions. As Andersen (1989) has observed, markedness has played an important role in linguistics for more than a century and while it has been explored from phonology to semantics, the focus here will be on markedness in syntax. According to the markedness principle, the *unmarked* member of a set of forms tends to be prototypical, more frequent, less presuppositionally loaded, and more widely distributed than the alternative *marked* form.

From a processing point of view, the question is whether markedness is anything more than a formal notion, or whether the distinction serves some role in language use. That is, what communicative functions are served by markedness? One possible answer is, of course, none at all. In this case, however, one would be hard-pressed to account for the ubiquity of markedness across languages. However, if markedness has psychological content, we should be able to discover the principles which govern the use of marked structures. Gundel, Houlihan & Sanders (1988), for example, have proposed that iconicity is a determining factor in markedness. Thus, for main clauses containing subordinate adverbial temporal clauses such as (2), it is the (a) member which Gundel et al. treat as unmarked.

- (2) a. After it snowed, classes were canceled.
- b. Classes were canceled after it snowed.

However, others (e.g., Bever 1970; Clark & Clark 1977; Givón 1983) have argued that it is clause order, not iconicity, which indicates markedness. Their claim is that in English the unmarked order is Main Clause (MC) followed by Subordinate Clause (SC) as in (2b), while SC + MC is the marked order. Givón (1987), in promoting a discourse basis for clause order, has suggested that preposed adverbial clauses in the marked (SC + MC) order tend to occur at major

thematic boundaries and depend for their interpretation on the preceding thematic material, whereas those with the unmarked MC + SC order tend to occur in the midst of a thematic unit². Giving this proposal some support, Fox (1987) found, in examining text data from Tagalog and English, that the marked member of a pair tended to be associated with the advent of a discourse unit (in her study, also a paragraph), while the unmarked member tended to be internal to the paragraph, a result also supported by Prideaux (1989). These combined results suggest that at least one function of marked structures might be found in a discourse management strategy whereby the speaker provides overt coding to signal shifts in discourse flow and content. However, since these results were all based on written text data, the question still remains as to whether such results are a function of the canons of the written language or if they reflect some more basic processing principles. The present study was undertaken to address this question. The hypothesis to be tested is:

MARKEDNESS HYPOTHESIS. Marked structures tend to be found initially in discourse units, while their unmarked counterparts tend to be found internal to the discourse units.

In this study, a short film clip taken from the movie *Adam's Rib* was used as stimulus. Forty participants (half male and half female) watched the clip and then later each provided an oral description of the scene to a friend (male or female) who had accompanied each participant but who had not seen the film. Participants were asked to provide enough detail so that the hearer, not having seen the film, could nevertheless have a clear idea of what had taken place. After seeing the film, each subject related the events in the film to the hearer, and with the permission of both participants, the description was taped. The tapes were then transcribed in conventional orthography, but with no editing.

The events within the clip constituted a self-contained unit, the massage scene, with Katherine Hepburn and Spencer Tracy playing husband and wife lawyers arguing opposite sides of a case. At home after a busy day in court, they are relaxing by giving each other massages while discussing the events of the day. Initially, Hepburn gives Tracy a backrub and when she is finished, she gently slaps his shoulder and says she is done. She then climbs on the table and Tracy gives her a massage. He turns the radio on, then off, since the song being played irritates him. She hums the song as he rubs her back, and then when he is finished, he gives her a hard slap on the bottom. This upsets her, they begin to argue, and she begins to cry. At this point, Tracy leaves the room, stands in the next room listening to her cry, and then returns. He bends over her to make up, and she kicks him in the shins and stomps out of the room. This ends the scene.

The film clip was subjected to a detailed examination by several analysts working first independently and then together, with the aim of constructing a fine-grained analysis of all the episodes and events contained in the scene. They concluded that the scene consists of nine major episodes, each containing (two to eight) independent but interrelated events. The nine major episodes of the scene are as follows (with each episode given a generic title and with the number of events within each episode indicated):

- E1. the woman massages the man (4)
- E2. they switch places (3)
- E3. the man massages the woman (8)
- E4. the woman gets upset (3)
- E5. they argue (2)
- E6. she cries (4)
- E7. he exits (2)
- E8. he returns (3)
- E9. she exits (3).

The event analysis provides an independent measure against which narrative descriptions can be assessed. It was anticipated that participants would not include everything in their descriptions but would rather focus on the major events. To assess the markedness hypothesis, structures consisting of a main and a subordinate clause were singled out for attention. Each such structure was coded for (a) whether it was unmarked (MC + SC) or marked (SC + MC) and (b) its location within an episode (i.e., initial or medial). An adverbial clause structure was scored as episode-initial only if it contained the first mention of any particular episode. The pooled data from the 40 participants are found in Table 5.

Structure	Episode Position	
	Initial	Medial
MC + SC (unmarked)	2	35
SC + MC (marked)	23	11

Table 5. Oral narrative adverbial clause distribution

Table 5 reveals that overall there was no significant difference in the total number of marked and unmarked structures used (37 unmarked versus 34 marked). However, the results reveal a strong tendency for the marked structures to appear initially in episodes and the unmarked structures medially, thereby

yielding strong support for the markedness hypothesis ($\chi^2_{(1)} = 30.089, p < .0001$). These oral data therefore provide strong empirical support for the hypothesis that one function of marked structures is to signal a new discourse unit.

An examination of some instances of the use of marked structures to redirect the narrative at episode boundaries proves interesting. In (3), the (female) narrator has just completed a discussion of the slapping episode and immediately turns to a discussion of the argument between the two characters.

- (3) *When he slapped her, they got into ... an argument about some court thing.*

A similar instance is found in (4), in which the (male) speaker uses two initial subordinate clauses to begin a new episode:

- (4) *and ... after they had started having their discussion about why he slapped her so hard, ... the heart of the matter came to the surface.*

In the same narrative, however, the speaker uses the unmarked MC+SC structure within an episode, maintaining the internal flow:

- (5) *He got kind of ... soft ... when he realized she was upset.*

In this instance, the narrator is discussing a change in the attitude of the male character toward his wife, and no redirection is involved at this point. These examples are characteristic of the distribution of the vast majority of the marked and unmarked structures in the oral narratives, all of which support the hypothesis that marked structures are used to code discourse boundaries.

A few counter-examples to the general tendency were found, the most interesting group of which is the eleven marked structures which appear internal to their respective episodes. A closer study of these reveals three important secondary functions of the marked structures: (a) they refer back to the first mention of the episode, (b) they signal the end of an episode, thereby heralding the onset of the next, or (c) they signal an especially important and salient event, a highlight of the entire film scene.

An example of the use of a marked structure within an episode unit to highlight a highly salient event can be found in (6).

- (6) *After he finished massaging her, he slapped her on the butt.*

Within the film, the actions of the man and woman slapping each other at the end of each massage serve as triggers for the development of the scene, which moves from a gentle to a more confrontational situation. The man's slapping his wife at the end of her massage is the major turning point of the scene. This

example also takes place at the end of an episode.

All eleven counter-examples of an internal marked structure fall into one of these three groups and, of these, three cases (including (6) above) serve both to highlight a salient event and also to mark the end of an episode. One common factor in all the uses of the marked structure seems to be that it serves to draw attention, either to a change in discourse unit or to an especially important event.

These oral data provide strong support for the hypothesis that marked structures cue discourse unit boundaries, as well as evidence of a highlighting function for the marked structures. The relevance of this study to our thesis is that a theory of markedness is not just an arbitrary formal exercise, but it does in fact have a functional basis. And as in the case of closure, markedness is simply not possible without a sophisticated concept of constituent structure.

4. Conclusions

The two studies reported here suggest that with the transition from a structural to a cognitive orientation, two concepts formulated by the structuralists, constituent structure analysis and markedness, are not simply notational devices, but rather they have cognitive content as well. Constituent structure provides us with a notation for representing linguistic structures and, coupled with closure, the two provide an explanation (*pace* Joos) of the empirical phenomena associated with the distributions of relative clauses in oral and written English.

Similarly, markedness is not just a convenient fiction to distinguish between basic and non-basic structures but has, in addition, important consequences for an understanding and explanation of one way that speakers of English manage information flow. Moreover, like closure, markedness could not even be formulated within the domain of syntax without a clear conception of constituent structure. We conclude that the structuralist notions of constituent structure and markedness are not just convenient fictions, but rather that each has empirical content in terms of language production and representation.

Acknowledgments

I dedicate this contribution to Professor Konrad Koerner, a scholar and friend who has contributed so much through both his own scholarly activities and his editorial skills to advance our science. This research was supported in part by SSHRC Research Grants 410-93-0109 and 410-96-0349. An earlier version of

some of the results discussed here was presented at the *Common Assumptions Conference* at the University of Chicago in April, 1997 and at two LACUS Conferences. I am grateful to Professor Victor Yngve for discussions dealing with foundational issues underlying much of the work reported here, although he by no means agrees with my position on the non-arbitrariness of constituent structure.

Author's address

Gary D. Prideaux
Department of Linguistics
University of Alberta
Edmonton, Alberta, CANADA T6G 2E7
Gary.Prideaux@UAlberta.ca

Notes

1. The studies are small parts of an extensive research program undertaken to investigate the contributions of cognitive and social factors to the organization of narrative discourse. I am grateful to the members of the Discourse Research Group at the University of Alberta for their extensive contributions over several years.
2. See Prideaux & Hogan 1993 for a more elaborate discussion of similar markedness results.

References

- Andersen, Henning. 1989. "Markedness Theory — The first 150 years". *Markedness in Synchrony and Diachrony* ed. by Olga M. Tomić, 11–46. Berlin & New York: Mouton de Gruyter.
- Bever, Thomas G. 1970. "The Cognitive Basis for Linguistic Structures". *Cognition and the Development of Language* ed. by John R. Hayes, 279–362. New York: John Wiley.
- Bloomfield, Leonard. 1933. *Language*. New York: Holt, Rinehart & Winston.
- Clark, Herbert H. & Eve V. Clark. 1977. *Psychology and Language*. New York: Harcourt Brace Jovanovich.
- Frazier, Lyn. 1985. "Syntactic Complexity". *Natural Language Parsing* ed. by David R. Dowty, Lauri Karttunen & Arnold M. Zwicky, 129–89. Cambridge: Cambridge University Press.
- & Janet Dean Fodor. 1978. "The Sausage Machine: A new two-part parsing model". *Cognition* 6.291–325.

- Gernsbacher, Morton Ann, ed. 1994. *Handbook of Psycholinguistics*. Toronto: Academic Press.
- Givón, Talmy, ed. 1983. *Topic Continuity in Discourse: Quantified cross-language studies. Typological studies in language, vol. 3*. Amsterdam & Philadelphia: John Benjamins.
- Givón, Tom, 1987. "Beyond Foreground and Background". *Coherence and Grounding in Discourse* ed. by Russell Tomlin, 173–88. Amsterdam & Philadelphia: John Benjamins.
- Gundel, Jeanette K., Kathleen Houlihan & Gerald Sanders. 1988. "On the Functions of Marked and Unmarked Terms". *Studies in Syntactic Typology. Typological studies in language*, 17 ed. by Mark Hammond, Edith A. Moravcsik & Jessica Werth, 285–301. Amsterdam & Philadelphia: John Benjamins.
- Harris, Zellig S. 1946. "From Morpheme to Utterance". *Language* 22.161–83.
- 1951. *Methods in Structural Linguistics*. Chicago: University of Chicago Press.
- Hockett, Charles F. 1948. "A Note on 'Structure'". *IJAL* 14.269–71.
- 1957. *A Course in Modern Linguistics*. New York: Macmillan.
- Joos, Martin 1950. "Description of Language Design". *Journal of the Acoustical Society of America* 22.701–8.
- Joos, Martin ed. 1958. *Readings in Linguistics*. Washington, D.C.: American Council of Learned Societies.
- Kimball, John P. 1973. "Seven Principles of Surface Structure Parsing in Natural Language". *Cognition* 2.15–47.
- Preston, W.D. 1948. Review of *Études Linguistiques Caribes II*, by C.H. de Goeje (Amsterdam, 1946). *IJAL* 14.31.
- Prideaux, Gary D. 1989. "Text Data as Evidence for Language Processing Principles: The grammar of ordered events". *Language Sciences* 11.27–42.
- & William J. Baker. 1986. *Strategies and Structures: The processing of relative clauses*. Amsterdam & Philadelphia: John Benjamins.
- & John T. Hogan. 1993. "Markedness as a Discourse Management Device: The role of alternative adverbial clause orders". *Word* 44:3.397–411.
- Sapir, Edward. 1925. "Sound Patterns in Language". *Language* 1.37–51.
- Twaddell, W. Freeman 1935. "On Defining the Phoneme". *Language Monograph* 16. Baltimore: The Linguistic Society of America.
- Wells, Rulon S. 1947. "Immediate Constituents". *Language* 23.81–117.

CHAPTER 26

From Ordered Rules to Ranked Constraints

John T. Jensen
University of Ottawa

1. Introduction

The conception of language as an instance of law-governed behavior has dominated linguistic research at least since the 19th C., when a major concern was discovering the laws of sound change in Indo-European and other language families. In the 20th C. the focus shifted to the study of the laws of synchronic grammars. In phonology, it was possible to formulate laws governing the distribution of allophones. Sounds that never contrast, and are phonetically similar, are grouped into a phoneme and rules state the environments in which each occurs. A more sophisticated application of the same technique allows for rules that account for morphophonemic alternations. If two sounds contrast in some environments but not in others, the noncontrasting environment is stated as a position of neutralization. For example, voiced and voiceless obstruents contrast in environments other than word-final position in German. In word-final position obstruents must be voiceless, which requires a rule of neutralization. As languages are analyzed in this way, numerous regularities of this type are stated, and the question arises as to how the rules relate to each other. As the analysis becomes more complex, it becomes apparent that not all rules can be true of all phonetic forms; thus, one rule may undo the effects of another or change its environment so that it no longer appears to have applied. This phenomenon, known as phonological opacity, is of crucial importance in developing appropriate models of phonological laws.

2. Ordered rules

Throughout most of the 20th Century the question of rule ordering has played a major role in phonological discussions. Chomsky & Halle (1968: 342) comment that "[t]he hypothesis that rules are ordered ... seems to us to be one of the best-supported assumptions of linguistic theory". Indeed, Bloomfield (1939) had already introduced ordered rules in his description of Menomini morphophonemics, though this device was not favored in pregenerative phonology, even in Bloomfield's own more theoretically oriented writings. Others take a much more dogmatic approach in denying the possibility of ordered rules on a priori grounds. For example Lamb (1966: 551) objects to phonological rules altogether, claiming that "[t]he process or rewriting formulation ... crept into morphophonemic description from diachronic linguistics". Lamb apparently regards rules as static relations between levels, a point to which we return. Kiparsky (1968: 121) rightly objects to Lamb's line of reasoning, pointing out that major proponents of process morphophonemics, such as Chomsky, Halle, Sapir, and Panini, were able to distinguish synchrony from diachrony. Harris (1951) reanalyzes one of Bloomfield's Menomini examples by stating the required rules so that they apply simultaneously. Although there are certain cases where this is possible, as we will see in section 4, there are many cases where no such reanalysis is possible.

As Chomsky & Halle emphasize, the issue is entirely empirical. They point out that it is possible to invent cases where simultaneously applied rules give the correct results, but where sequentially applied rules do not. Since the real languages that have been investigated in sufficient detail to test the competing hypotheses of rule application support the hypothesis that rules apply in sequence (even though some may be consistent with simultaneous application), the sequential ordering hypothesis is better supported empirically. Chomsky & Halle introduce the term "depth of ordering" to describe the situation. The depth of ordering of a grammatical description is the number of sequentially ordered rules in a subsequence of rules "with the property that the grammar becomes more complex if any two successive rules of this subsequence are interchanged in the ordering" (Chomsky & Halle 1968: 18, fn. 4). According to Bever (1967), the depth of ordering in Bloomfield (1939) is at least eleven, and Chomsky (1951) demonstrates a depth of ordering of twenty-five in Modern Hebrew. Nevertheless, the question of rule ordering quickly became a very controversial one, as we shall see.

In order to have some concrete basis for discussion, let us review one of the best known examples of a phonological analysis with ordered rules, Kenstowicz

& Kisseberth's (1979) analysis of Yawelmani. This analysis can then serve as a basis for discussion of later approaches. At the heart of the analysis are the four rules of Epenthesis, Vowel Harmony, Lowering, and Shortening, which must apply in that order. Epenthesis inserts the vowel *i* between two consonants if a third consonant or a word boundary follows, as in (1).

$$(1) \quad \emptyset \rightarrow i / C_C \left\{ \begin{array}{c} C \\ \# \end{array} \right\}$$

Vowel Harmony makes a vowel round and back following a round vowel that has the same value for the feature [high], as in (2).

$$(2) \quad \left[\begin{array}{c} +\text{syll} \\ \alpha\text{high} \end{array} \right] \rightarrow \left[\begin{array}{c} +\text{round} \\ +\text{back} \end{array} \right] / \left[\begin{array}{c} +\text{syll} \\ +\text{round} \\ \alpha\text{high} \end{array} \right] C_0_ \text{ (left-to-right iterative)}$$

Epenthesis must apply before Vowel Harmony. Epenthesis feeds Vowel Harmony if the epenthetic vowel follows a high round vowel, as in the derivation in (3).

- (3) /ʔugn + hin/ underlying representation
 ʔuginhin Epenthesis
 ʔugunhin Vowel Harmony (first iteration)
 [ʔugunhun] Vowel Harmony (second iteration); phonetic representation
 "drinks"

Epenthesis bleeds Vowel Harmony if the inserted vowel appears between two nonhigh vowels, the first of which is round, as in the derivation (4).

- (4) /logw + xa/ underlying representation
 logiwxa Epenthesis
 — Vowel Harmony (blocked)
 [logiwxa] phonetic representation
 "let us pulverize"

Since the vowels *o* and *a* are both [−high] in (4), Vowel Harmony would have turned *a* to *o* had not Epenthesis applied first. The result would have been the ungrammatical representation **logiwxo*. Both surface forms produced in (3) and (4) are transparent, however, in the sense that the phonetic representation contains the conditions for Vowel Harmony when it has applied, as in (3), or excludes the conditions for Vowel Harmony when it has not applied, as in (4).

redundant in certain cases but not others. For example [round] is distinctive for back vowels in English but not for front vowels. The question then becomes which feature to specify in underlying representations. It is generally assumed that [back] is specified for all vowels while [round] is specified only for back vowels.

In addition to such constraints on underlying representations, Kisseberth (1970) introduced a category of *derivational constraints*. These allow for the simplification of a number of phonological rules that participate in a "conspiracy" to produce a certain kind of output. Kisseberth discusses a number of rules of Yawelmani that seem to conspire to ensure that phonetic representations do not contain three consecutive consonants or two consecutive consonants at the beginning or end of a word. The Epenthesis rule (1) is one such rule; another is a rule that deletes a short vowel in the context VC____CV. Kisseberth argues that the latter rule can be simplified to the context C____C if the rule is subject to the derivational constraint prohibiting CCC. A number of problems arise in the formalization of MSCs and derivational constraints, a detailed discussion of which would take us too far afield. One specific problem is that the proposal to allow blank specifications in underlying representations gives rise to the possibility of interpreting the blank as a third value, distinct from both + and -. In an important review of these questions, Stanley (1967) concluded that underlying representations could not contain blank specifications, even for redundancies, just because of this problem.

The development of lexical phonology, especially Kiparsky (1982), changes the picture dramatically. Kiparsky argues that MSCs are not properly stated on *morphemes*, but rather on *words*, more specifically on *lexical items* as this term is defined in lexical phonology. This is certainly true of syllabification rules. Many languages contain morphemes that cannot form syllables on their own, say the plural suffix /-z/ of English, but are syllabifiable only in conjunction with a noun stem. Kiparsky reviews four specific problems with MSCs: the question of whether a generalization is properly stated as a condition or a rule, the question of why conditions often duplicate rules, the domain on which constraints are stated (the morpheme or the word) and finally the problem of the level where the constraint holds: underlying or derived. Kiparsky shows that MSCs are actually rules of the lexical phonology. This solves the question of whether the generalizations are constraints or rules: in Kiparsky's treatment they are all rules. The duplication problem is solved in the same way. The domain problem is resolved by the principles of the theory as well. Rules are stated as applying on a particular stratum of the lexicon and so apply within the domain of words formed at that stratum. The level problem is reduced to the general property of phonological rules in generative phonology: a rule applies at the stage in the

derivation where it fits in the overall ordering, here expanded to include the stratum of the lexicon where the rule applies. This organization allows the solution of an additional question, that of strict cyclicity. Kiparsky adopts a form of radical underspecification, which we can give as (8).

- (8) For any feature *F*, only one specified value [+*F*] or [−*F*] may appear in a given environment in underlying representations.

Kiparsky also proposes to state the Strict Cycle Condition (SCC) as in (9).¹

- (9) Cyclic rules can change structure only in derived environments.

Kiparsky illustrates the effects of (9) with the English rule of Trisyllabic Laxing (TSL). There are four cases to consider, which we list in (10).

- (10) a. Structure-changing in derived environment: *sanity* (compare *sane*)
 b. Structure-filling in nonderived environment: *sycamore*
 c. Nonstructure-changing in nonderived environment: *nightingale*
 d. Lexical exception: nonstructure changing in derived environment: *obesity*

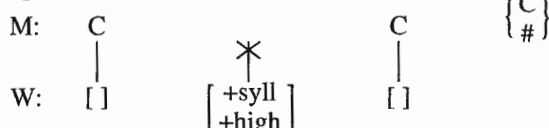
The normal case of the operation of this rule is (10a). The morpheme *sane* has an underlying tense vowel. The addition of the suffix *-ity* provides a derived trisyllabic environment, and the rule applies. Words like *sycamore*, (10b), would be obliged to have an underlyingly specified [−tense] vowel by the full specification argued for by Stanley. Kiparsky argues that TSL can indeed apply in nonderived forms. The first vowel of *sycamore* can be unspecified for [tense] in underlying representation and TSL applies in a structure-filling capacity here, since (9) only restricts the structure-changing function of cyclic rules. This neatly solves the problem that Mascaró (1976) raised in an earlier statement of the SCC. Mascaró disallowed all application of cyclic rules in nonderived environments, thus being forced into the unfortunate position that stress had to be specified in underlying representations, even though it is largely predictable.

Words like (10c), *nightingale*, caused considerable difficulty for the *SPE* version of TSL. *SPE* makes the rather improbable proposal that *nightingale* is derived from an underlying representation whose first syllable contains a lax [ɪ] followed by a velar fricative [x]. TSL then has no effect on this vowel. Later rules convert the sequence [ɪx] to [i] (tense *i*) which ultimately becomes the diphthong [ay] by Vowel Shift and other rules. Kiparsky (1982) points out that this solution is not available for the majority of words in this class, such as *stevedore*. While words like *nightingale* and *stevedore* seem somewhat exceptional

compared to examples like *sycamore* (monomorphemic words that conform to TSL), they should not have to be marked as lexical exceptions in the same way as *obese*. Kiparsky proposes simply that such words have lexically marked [+tense] vowels in their first syllables, which then cannot undergo TSL because of (9). This makes them slightly more marked than words like *sycamore*, whose first vowels are unspecified for [tense]. Only words like *obese* have to be marked as true lexical exceptions, since they fail to undergo TSL even in derived environments (*obesity*).

4. The end of rules?

Goldsmith considers that the theory of lexical phonology, summarized in the last section, “depends heavily on what appear to be thoroughgoing uses of an implausible metaphor involving space and time...” (1993b: 21). Goldsmith proposes in its stead a theory of “harmonic phonology”. In the same volume Lakoff (1993) proposes a theory of “cognitive phonology”. These two approaches are actually rather similar and can be discussed together. Goldsmith accepts Kiparsky’s proposal that lexical rules and MSCs are essentially the same thing, but emphasizes the phonotactic function over the rule function. Goldsmith also reintroduces stratificationalism (Lamb 1966) and even repeats Lamb’s objections to derivations as somehow unwarranted intrusion of historical linguistics into synchronic phonology, despite Kiparsky’s (1968) rejoinder. The system of harmonic phonology proposed by Goldsmith and Lakoff posits three levels, which they call the M-level (morphophonemic level), the W-level (word level), and the P-level (phonetic level). On the M-level, morphemes are phonologically specified. On the W-level, expressions are structured into well-formed words and syllables with a minimum of redundant phonological information, and the P-level is a broad phonetic description that interfaces with articulatory and acoustic devices. The operation of this type of grammar can best be seen by summarizing Goldsmith & Lakoff’s analysis of Yawelmani, to compare with our earlier generative discussion (section 1). Goldsmith and Lakoff assume two types of “constructions” (Lakoff’s term) instead of rules. Intralevel constructions state well-formedness constraints within a level. Interlevel constructions express correlations across levels. Interlevel constructions are direction neutral and the constraints expressed by all constructions are satisfied simultaneously. There are also universal or language-specific defaults and default identity is assumed between levels, where not overridden by an interlevel construction. Lakoff analyzes Yawelmani Epenthesis as an M-W construction, which he formalizes as in (11).

(11) *Epenthesis*

And he assumes a markedness default value for high vowels as in (12).

(12) *High Vowel Markedness*

If $\left[\begin{smallmatrix} +\text{syll} \\ +\text{high} \end{smallmatrix} \right]$ then $[-\text{back}]$

Lakoff analyzes Vowel Harmony as a W-construction, as in (13).

(13) *Vowel Harmony*

W: if $\left[\begin{smallmatrix} +\text{syll} \\ +\text{round} \\ \alpha\text{high} \end{smallmatrix} \right]$ C_0X , then if $X = \left[\begin{smallmatrix} +\text{syll} \\ \alpha\text{high} \end{smallmatrix} \right]$ then $X = \left[\begin{smallmatrix} +\text{round} \\ +\text{back} \end{smallmatrix} \right]$

Shortening and Lowering are W-P constructions in this model. Lakoff formulates Shortening as in (14) and Lowering as in (15).

(14) *Shortening*

W:	$\left[\begin{smallmatrix} +\text{syll} \\ +\text{high} \end{smallmatrix} \right]$	C	C
P:	$[-\text{long}]$		

(15) *Lowering*

W:	$\left[\begin{smallmatrix} +\text{syll} \\ +\text{long} \end{smallmatrix} \right]$
P:	$[-\text{high}]$

It is not clear why this analysis is less an “implausible metaphor” than lexical phonology. Despite the change in terminology, and Lakoff’s claim that cognitive phonology “drastically changes what can and cannot be done in phonology” (1993: 145), Lakoff’s analysis has in effect incorporated the same ordered rules

as the generative analysis given in section 2. One thing that makes this possible is that the generative analysis could allow Shortening and Lowering to apply simultaneously rather than in that order (although applying them in the opposite order obviously gives incorrect results). Generative phonology frequently noted that certain rule interactions could be obtained by simultaneous application, but that other cases required sequential application. However, all the cases that could be done using simultaneous application were also obtained by sequential application, and so the model opted for a single mode of ordering rather than two. Therefore, it seems accidental that cognitive phonology is able successfully to reformulate generative analyses involving relatively few rules. Neither Lakoff nor Goldsmith discuss how this model would account for strict cyclicity effects (section 2), nor do they extend it to really complex cases. The lexical phonology of English requires at least twenty-two rules on stratum 2 (the word level) alone (Jensen 1993), with many crucial orderings and complex interactions, which would be difficult to account for in this framework.

The cognitive model did not develop beyond these preliminary efforts. It was overshadowed by the rise of Optimality theory (section 5), which does try to deal with cases of opacity and strict cyclicity, though with limited success to date. To close this section, we should emphasize that Lexical Phonology was not intended as a processing model, as the following quote from Kiparsky makes clear.

This does *not* imply that the speaker or hearer need in any way mentally "derive" the words he says or hears by means of such rules as Velar Softening. What it *does* mean is that the alternations they govern belong to the regular phonological patterns of English, while for example a hypothetical *k-s* alternation in the reverse context, such as **criti[k]ize ~ *criti[s]al* would be irregular. The claim made is that someone who knows English implicitly knows that pattern, and will under appropriate circumstances recognize the difference between regular and irregular alternations, though he may not be able, even after reflection, to verbalize the rules that underlie it. (Kiparsky 1982: 34)

5. Optimality theory

Optimality theory (OT) has strong roots in the tradition of phonological rules as well as in the harmonic/cognitive approaches to phonology. While OT bills itself as a radically new approach to phonology, it in fact incorporates many elements of standard generative phonology. It retains the notion of phonology as a device relating input to outputs, along with the relatively abstract underlying representations (of Kenstowicz & Kisseberth's analysis of Yawelmani, for example). It

differs from standard generative phonology in placing the entire burden of determining the grammatical output on a system of constraints, which are universal, but whose ranking is a matter of language-specific grammars. In place of generative rules, OT allows a very unrestricted component called GEN to produce a large (perhaps infinite) number of candidate outputs. An OT analysis may indeed look very different from a generative analysis (or from the cognitive/harmonic phonology analysis, for that matter). To account for the phenomenon of Epenthesis in Yawelmani, introduced in section 2, OT would invoke a constraint *Complex, which prohibits more than one consonant in an onset or coda. This constraint is undominated in Yawelmani, which ensures that it is unviolated. GEN supplies candidate outputs that conform to *Complex by virtue of including epenthetic vowels. In the earliest versions of OT, epenthetic elements are indicated by a box, \square , suggesting an empty syllabic position whose content is determined by phonetic interpretation. In fact, GEN supplies candidates with many epenthetic positions. Clearly, we do not wish to allow epenthesis all over the place — we need it just in the right places. Each epenthetic position constitutes a violation of the constraint Fill, which requires that syllabic positions be filled by material from the input form. This is a case of constraint violation. Given an input /logw + hin/, an output with no epenthesis violates *Complex. An output with epenthesis violates Fill. The dominant position of *Complex ensures that it will never be violated, so in this case Fill must be violated by the optimal output form. But we don't want to violate it more than necessary: one epenthetic vowel is sufficient to ensure conformity to *Complex. Epenthesis in the correct position is ensured by the constraint Align-Right, defined by McCarthy & Prince (1993: 116) as in (16).

(16) *Align Right*

Align(Stem, R, σ , R)

"For every stem, there exists a syllable such that the right edge of that syllable is aligned with the right edge of the stem"

The results are displayed in a tableau such as (17). The tableau lists a few of the plausible candidates and a selection of the universal constraints involved in the selection of the optimal candidate, which is picked out by the symbol \Rightarrow .

(17)

/logw + hin/ 'pulverizes'	*Complex	Fill	Align Right
lo.g□w.hin		*	
logw.hin	*!		
log.whin	*!		
log.w□.hin		*	*!
log□.w□.hin		**!	*

In each candidate the syllable boundaries are indicated by dots. The first candidate is selected as optimal, even though it violates Fill, because other possible candidates violate *Complex, which is higher ranked. The second and third candidates each violate *Complex — the second by having a complex coda, the third by having a complex onset. The fourth candidate is equal to the first in terms of the first two constraints, but violates Align Right, which the first candidate conforms to. The fifth candidate violates Fill twice (and also Align Right). The violation that eliminates a candidate from the competition is marked by an exclamation point. Cells are shaded if the constraint the cell falls under plays no role in the selection of a candidate. Thus the cells for Fill and Align Right are shaded for the second and third candidates, which have been eliminated by *Complex, and the cell under Align Right is shaded for the fifth candidate because the crucial violation for this candidate is the second violation of Fill.

The main virtue claimed for OT is that its constraints can be violated, if that violation is needed to select a candidate that better conforms to a higher ranked constraint. In Yawelmani, the constraint Align Right is violated by some optimal phonetic forms. To see this, consider a tableau for /logw + ol/. No epenthesis is required to conform to *Complex, so epenthetic candidates will be eliminated by Fill. To conform to *Complex, the successful candidate must be syllabified correctly, i.e., between the *g* and the *w*. Candidates that syllabify before or after this cluster violate *Complex. The successful candidate is the first one in (18), which conforms to both *Complex and Fill, but violates Align Right.

(18)	logw + ol 'might pulverize'	*Complex	Fill	Align Right
	log.wol			*
	lo.gwol	*!		
	lo.g□.wol		*!	*
	logw.ol	*!		

Each of the constraints in tableaux (17) and (18) illustrates a different type of constraint. *Complex is a *structural* constraint that specifies the types of structure permitted in phonetic forms; others of this kind are Onset, which requires syllables to have onsets, and NoCoda, which disallows codas. NoCoda is fairly low ranked in Yawelmani, since syllable codas are quite common, but Onset is ranked high. Fill is a *faithfulness* constraint. These are constraints that ensure that outputs resemble their corresponding inputs as much as possible. Another faithfulness constraint is Parse, which holds that underlying material must be parsed into syllable structure. Align Right is an interface constraint that regulates the relations between morphological structure and phonological structure. This early version of OT assumes that GEN is regulated by *containment*, which says that all candidate outputs contain the input; that is, there is no literal deletion of elements, though some input segments may remain unparsed and so unrealized phonetically. This approach allows all constraints to be stated on outputs. In later versions containment is rejected, and deletion is allowed. In this version, faithfulness is stated in terms of correspondence constraints, and constraints referring to inputs and outputs are used, much like the cross-level constructions of cognitive phonology.

A major success of OT lies in its treatment of prosodic morphology, in particular infixation and reduplication. As an example, consider Tagalog infixation. The prefix *um-* is fully word initial before a vowel-initial stem, but is infixal after a consonant or consonant cluster, as shown in the data of (19).

(19)	aral	umaral	"teach"	
	sulat	sumulat	"write"	*umsulat
	gradwet	grumadwet	"graduate"	*umgradwet, *gumradwet

McCarthy & Prince (1993: 103) demonstrate that this follows if the structural constraint NoCoda dominates the morphological constraint that aligns *um-* to the left of a stem. They refer to this latter constraint as Align-*um*, but it is really an instance of a more general constraint that demands left alignment of prefixes.

Align-*um* is an unlikely candidate for a *universal* constraint.

- (20) Align-*um*: Align ([*um*]_{AF}, L, Stem, L)

(20)	{ <i>um</i> , <i>gradwet</i> }	NoCoda	Align- <i>um</i>
a.	<i>um.grad.wet</i>	***!	
b.	<i>gum.rad.wet</i>	***!	g
c.	<i>gru.mad.wet</i>	**	gr
d.	<i>grad.wu.met</i>	**	gra!dw

In this tableau, the input is seen as an unordered set of prefix plus stem. The optimal output is the candidate with the fewest violations of NoCoda that is consistent with leftmost positioning of the prefix. Since NoCoda dominates Align-*um*, candidates with more than two NoCoda violations are excluded, since candidates are available with only two. Of the last two candidates, candidate (20c) has the best positioning of *um*- and so is selected as optimal.

Once faithfulness was expressed in terms of correspondences (rather than Parse and Fill on outputs alone), other correspondences were discovered. For example, reduplication is expressed as a correspondence between the base and the reduplicant. Correspondences between several outputs were proposed (Benua 1995) to account for truncation phenomena, such as the truncation of the name *Larry* to *Lar*, with [æ], even though this vowel is normally disallowed in monosyllables ending in *r*.

We illustrated the basic principles of OT using data from Yawelmani, but without considering Vowel Harmony. Yawelmani Vowel Harmony is a problem for OT because its conditions are not uniformly met in either underlying (input) or surface (output) forms. As we saw in section 2, underlying high vowels that are lowered count as high for harmony, but epenthetic vowels (which are not present underlyingly) also count as high and may interrupt a nonhigh harmony domain. In cognitive phonology the intermediate W-level serves as the domain for Yawelmani Vowel Harmony, but no such intermediate level is allowed in OT. Cole & Kisseberth (1995) try to account for this in their version of OT, which they call Optimal Domains Theory (ODT). Their height requirement on Yawelmani Vowel Harmony is stated as a uniformity constraint, which we give as (21).²

- (21) Uniformity: The harmony domain must be monotonic: High or Low. If conflicting High and Low domains are present, the faithful height domain is counted.

This formulation makes reference to the input value of [high] if it is present and differs from the output; otherwise it refers to the output value. This works but seems a rather roundabout way of getting what cognitive phonology achieves directly with its W-level.

The Yawelmani case, along with other opaque interactions, is examined in a recent development of OT known as sympathy theory (McCarthy 1997; Itô & Mester 1997). In a nutshell, this theory seeks to account for opaque interactions by invoking a constraint that requires faithfulness to a specified failed candidate. The proposal can be illustrated with the tableau in (22).

(22)

		/deʃʔ/	Coda Cond	*Max- V _ʃ	Max- C _{IO}	Dep-V _{IO}
opaque	a.	deʃe			*	i*!
transparent	b.	deʃ		*!	*	
sympathetic	c.	deʃeʔ	*!		✓	*

The Coda Condition prohibits certain consonants, including [ʔ], from appearing in the coda. Max-C_{IO} is a faithfulness constraint, similar to Parse in the earlier version of the theory, that requires all consonants in the input to appear in the output. Dep-V_{IO} is another faithfulness constraint, like earlier Fill, that requires all output vowels to have a source in the input. If we stick with normal constraints for the moment, disregarding the sympathy constraint *Max-V_ʃ, we would expect the transparent candidate (22b) to be the actual output. To select the actual (opaque) output (22a), McCarthy designates a specific faithfulness constraint, in this case Max-C_{IO}, as the “selector” constraint. The sympathetic candidate (22c) is the candidate that best satisfies the constraint hierarchy modified so that the selector is top ranked. Of the candidates in (22), only (22c) satisfies this constraint, indicated by the ✓. The candidate so selected is designated the sympathy candidate, marked by the symbol *. McCarthy then postulates a faithfulness constraint, *Max-V_ʃ, demanding that all vowels in the sympathy candidate appear in the output. If this sympathy constraint is appropriately ranked, the correct output is selected by the hierarchy. Note that the sympathy faithfulness constraint is itself dominated by the Coda Condition; otherwise, the sympathetic candidate would incorrectly be selected as the actual output.

While this works, it achieves the right result in a rather roundabout way, compared to the way this would be done using ordered rules. The derivation in (23) illustrates this.

- (23) /deš?/ underlying representation
 deše? Epenthesis
 {deše} ?-Deletion; phonetic representation

In fact, the sympathy candidate in tableau (22) is precisely the intermediate representation in derivation (23). McCarthy argues that this is not necessarily always the case, and considers how the Yawelmani case would be analyzed in this approach, using two distinct sympathetic candidates. The number of relations has exploded from the small number (input-output, base-reduplicant, output-output, and a few others) to an infinite number, since each of the infinite number of candidates for any given form is in a relation to every other candidate. As the depth of ordering of a derivational analysis increases, so will the number of sympathy candidates required in this type of analysis. McCarthy claims that sympathy theory cannot duplicate the full expressive power of derivational theory, and in particular that derivations in which one rule undoes the effects of another cannot be produced in a sympathy theoretic analysis. This result is not clear either, if the class of possible selector constraints is expanded beyond faithfulness constraints, as proposed by Itô & Mester 1997. However, the development of sympathy theory clearly points the way to further research in properties of serial derivations as well as optimality theory leading to a better understanding of these issues.

Author's address

John T. Jensen
 Department of Linguistics
 University of Ottawa
 Ottawa, Ontario, Canada K1N 6N5
 jjensen@aix1.uottawa.ca

Notes

1. Kiparsky (1982) claims that the Strict Cycle Condition is actually derived from the Elsewhere Condition, and need not be stipulated in the grammar. Kiparsky (1985) gives a somewhat more formal statement of the SCC that basically has the effect of (9). For the purpose of the discussion here, the statement in (9) is sufficient.
2. The term "Low" in this formulation should be understood as meaning [-high]. Their analysis assumes privative features, which is a question entirely independent of the issue under discussion here.

References

- Benua, Laura. 1995. "Identity Effects in Morphological Truncation". *University of Massachusetts Occasional Papers in Linguistics* 18 (Papers in Optimality Theory) ed. by Jill Beckman, Laura Walsh-Dickey & Suzanne Urbanczyk, 77–136. Amherst: GLSA.
- Bever, Thomas G. 1967. *Leonard Bloomfield and the Phonology of the Menomini Language*. Ph.D. dissertation, MIT.
- Bloomfield, Leonard. 1939. "Menomini Morphophonemics". *Travaux du Cercle linguistique de Prague* 8.105–115. [Reprinted in Makkai 1972. 58–64.]
- Chomsky, Noam. 1951. *Morphophonemics of Modern Hebrew*. M.A. thesis, University of Pennsylvania.
- & Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper & Row.
- Cole, Jennifer S. & Charles W. Kisseberth. 1995. "Restricting Multi-Level Constraint Evaluation: Opaque rule interaction in Yawelmani vowel harmony", Cognitive Science Technical Report UIUC-BI-CS-95–06. The University of Illinois, Urbana-Champaign.
- Goldsmith, John, ed. 1993a. *The Last Phonological Rule*. Chicago: University of Chicago Press.
- 1993b. "Harmonic Phonology". Goldsmith 1993a. 21–60.
- Harris, Zellig S. 1951. *Structural Linguistics*. Chicago: University of Chicago Press.
- Itô, Junko & Ralf-Armin Mester. 1997. "Sympathy Theory and German Truncations". Rutgers Optimality Archive.
- Jensen, John T. 1993. *English Phonology*. Amsterdam: Benjamins.
- Kenstowicz, Michael J. & Charles W. Kisseberth. 1979. *Generative Phonology: Theory and description*. New York: Academic Press.
- Kiparsky, Paul. 1968. "How Abstract is Phonology?" *Explanation in Phonology*, ed. by Paul Kiparsky, 119–163. Dordrecht: Foris.
- 1982. "Lexical Morphology and Phonology". *Linguistics in the Morning Calm* ed. by I. S. Yang, 3–91. Seoul: Hanshin.
- 1985. "Some Consequences of Lexical Phonology". *Phonology Yearbook* 2.85–138.
- Kisseberth, Charles W. 1970. "On the Functional Unity of Phonological Rules". *Linguistic Inquiry* 1.291–306.
- Lakoff, George. 1993. "Cognitive Phonology". Goldsmith 1993a. 117–145.
- Lamb, Sydney M. 1966. "Prolegomena to a Theory of Phonology". *Language* 42.536–573. [Reprinted in Makkai (1972), 606–633.]
- Makkai, Valerie Becker, ed. 1972. *Phonological Theory: Evolution and current practice*. New York: Holt, Rinehart & Winston.
- Mascaró, Joan. 1976. *Catalan Phonology and the Phonological Cycle*. Ph.D. dissertation, MIT.

- McCarthy, John J. & Alan S. Prince. 1993. "Generalized Alignment". *Yearbook of Morphology 1993*. 79–153.
- McCarthy, John J. 1997. "Sympathy and Phonological Opacity". Rutgers Optimality Archive.
- Stanley, Richard. 1967. "Redundancy Rules in Phonology". *Language* 43.393–436.

PART V

Indo-European Linguistics

CHAPTER 27

Next of Kin

The search for relatives of Indo-European

Allan R. Bomhard
Charleston, South Carolina

From the very earliest days of Indo-European comparative linguistics, there have been speculations about the possible genetic relationship of Indo-European to other language families. Though, in the course of study, many striking similarities were noted between Indo-European and certain other language phyla, notably Uralic and Afrasian (formerly called Hamito-Semitic, Semito-Hamitic, Afro-asiatic, Erythraic, and Lisramic), truly convincing evidence of distant linguistic relationship was simply not brought forth. Indeed, much of the early work was not of high quality and did more to discredit the attempt to discover possible relatives of Indo-European than to help. Gradually, the intellectual climate, especially in the United States, became hostile to long-range comparison, and Indo-European remained an orphan with no known relatives.

In the first half of the last century, no less a figure than one of the founders of Indo-European comparative grammar, Franz Bopp, investigated possible relationship of Indo-European with Kartvelian (in 1846 and 1847) on the one hand and with Malayo-Polynesian (in 1840) on the other. In the mid-1860s, Rudolf von Raumer (in 1863) and Graziadio Ascoli (in 1864) claimed that Indo-European and Semitic were related. At about the same time (in 1869), Vilhelm Thomsen proposed relationship between Indo-European and Finno-Ugrian. This proposal was later (in 1879) explored in depth by the Estonian Nicolai Anderson and (in 1900) by the British phonetician Henry Sweet. Unfortunately, Anderson's work contained too many errors to be of lasting value. However, insightful and solid contributions were made concerning the possible relationship of Indo-European and Uralic during the current century by the Swedish Uralicist Björn Collinder. Towards the end of the last century (1873), the Semiticist Friedrich Delitzsch investigated lexical parallels between Indo-European and Semitic.

Then, at the beginning of this century, the Danish linguist Hermann Möller, in the course of several publications, attempted to show that Indo-European and Semitic might be related. Möller's work was later continued by the French linguist Albert Cuny, whose last publications date from the mid-1940s. Möller's and Cuny's efforts were generally not highly regarded by the scholarly community. One exception was Möller's student Holger Pedersen, who not only coined the term "Nostratic" but who also expanded the definition to include Indo-European, Semitic, Samoyed and Finno-Ugrian, Turkish, Mongolian, Manchu, Yukaghir, and Eskimo (cf. Pedersen 1931: 337–338). In 1969, Linus Brunner published a detailed comparison of the Indo-European and Semitic vocabularies, and this was followed in 1980 by a wider comparison of languages undertaken by Kalevi E. Koskinen. We should note also that, though the investigation of problems relating to distant linguistic comparison was generally ignored by the vast majority of mainstream linguists, the field was never completely dormant — a small but persistent group of scholars (Pentti Aalto, Knut Bergsland, Václav Blažek, René Bonnerjea, Karl Bouda, Bojan Čop, Heinz Fähnrich, Carleton T. Hodge, G. A. Klimov, D. H. Koppelman, Saul Levin, Karl Menges, Roy Andrew Miller, Mikolas Palmaitis, Stephen A. Tyler, Ants-Michael Uesson, C. C. Uhlenbeck, to name but a few of the many scholars working on long-range comparison) has continued to work, throughout the better part of this century, on binary (or, in rare cases, wider) comparisons of various languages that are currently considered to belong to the Nostratic macrofamily. For comprehensive bibliographies listing publications dealing with distant linguistic comparison, cf. Hegedűs 1992, Landsberg 1986, and Bomhard & Kerns (1994: 715–864).

Recently, beginning in the mid-1960s, the intellectual climate has slowly begun to turn around, and a growing number of linguists, especially in the former Soviet Union, have begun to turn attention toward investigating distant linguistic relationship. The revived interest was sparked by the work of Vladislav M. Illič-Svityč and Aaron B. Dolgopolsky, who first started working independently and, at a later date, through the efforts of their mutual friend Vladimir Dybo, cooperatively. Their work, though not without its own shortcomings, was the first successful demonstration that certain language phyla of northern and central Eurasia, the Indian subcontinent, and the ancient Near East might be genetically related. Following a proposal made in 1903 by Holger Pedersen, they employed the name "Nostratic" to designate this grouping of languages. In particular, Illič-Svityč, in the course of several publications, culminating in his posthumous comparative Nostratic dictionary, which is still in the process of publication, included Indo-European, Kartvelian, Afasian, Uralic, Dravidian, and Altaic in his version of the Nostratic macrofamily. From his very earliest

writings, Dolgopolsky also included Chukchi-Kamchatkan and Eskimo-Aleut.

Before his tragic death in an automobile accident on 21 August 1966, Illič-Svityč had planned to prepare a comparative Nostratic dictionary listing over 600 Nostratic roots and tracing their development in detail in each of the daughter languages in which they were attested. He had published a preliminary report on his work in 1965 entitled (in English translation) "Materials for a Comparative Dictionary of the Nostratic Languages (Indo-European, Altaic, Uralic, Dravidian, Kartvelian, Hamito-Semitic)". Working diligently, literally devoting all of his energy to the project, he had managed to prepare the entries for approximately 350 roots. After his death, Illič-Svityč's work was prepared for publication by the dedicated efforts of Rimma Bulatova, Vladimir Dybo, and Aaron Dolgopolsky, with the result that the first volume of the dictionary appeared in 1971, containing 245 entries. A second, smaller volume appeared in 1976, listing entries 246 through 353 and ending with an index — this completed all of the material prepared by Illič-Svityč himself (by the time this volume appeared, Dolgopolsky was in the process of emigrating to Israel). Finally, the first fascicle of volume three appeared in 1984, containing entries 354 through 378, none of which was prepared by Illič-Svityč — it represents the collective efforts of a team of scholars.

In the meantime, Dolgopolsky has continued to make important contributions to Nostratic studies, especially in a 1984 paper on Nostratic pronouns, and currently has material to support the reconstruction of just over 2,000 Nostratic roots. Other Russian scholars have also done important research into problems affecting Nostratic — mention should be made of the work of Alexandra Y. Aikhenvald, N. D. Andrejev, M. S. Andronov, Vladimir Dybo, Eugene Helimskij, Vjačeslav V. Ivanov, G. Kornilov, Oleg Mudrak, Vitaly V. Shevoroshkin, Sergej A. Starostin, V. A. Terentjev, Vladimir N. Toporov, and V. L. Tsyburskij, among others. Though not Russian (but clearly someone who belongs to the Moscow School), special recognition must be given to the Czech scholar Václav Blažek, who has published many important papers, most of which deal with the common Nostratic lexicon. Another who should be noted is Alexis Manaster Ramer — he has been publishing a number of interesting papers on Nostratic recently.

Beginning with an article that appeared in *Orbis* in 1975, Allan R. Bomhard published several studies, culminating in a 1984 book entitled *Toward Proto-Nostratic: A new approach to the comparison of Proto-Indo-European and Proto-Afroasiatic*, in which he tried to show that Indo-European and Semitic (later expanded to include all of Afrasian) might be distantly related. Reviews of this book as well as discussions with colleagues prompted Bomhard to expand the scope of his research to include other language families. This resulted in the

publication in April 1994 by Bomhard of a joint monograph with John C. Kerns entitled *The Nostratic Macrofamily: A study in distant linguistic relationship*. It was Kerns who prepared the chapter dealing with Nostratic morphology. This book supplies a great deal of lexical evidence from the Nostratic daughter languages to support the reconstruction of 601 Proto-Nostratic roots. In an article published in *Orbis* in 1995, Bomhard supplies material to support an additional 29 Proto-Nostratic roots. Bomhard continues to gather lexical data and plans future articles listing still more common Nostratic roots — 21 new etymologies were proposed in his 1996 book entitled *Indo-European and the Nostratic Hypothesis*. It should be noted that Bomhard's views on Nostratic differ somewhat from those of Illič-Svityč (and others who follow his system).

Joseph Greenberg is currently preparing a two-volume work to be entitled *Indo-European and its Closest Relatives: The Eurasiatic language family*. The first volume will deal with grammar, and the second will deal with lexicon. Greenberg includes Indo-European, Uralic-Yukaghir, Altaic (Mongolian, Chuvash-Turkic, and Manchu-Tungus), Japanese-Korean (Korean, Ainu, and Japanese-Ryukyuan), Gilyak, Chukchi-Kamchatkan, and Eskimo-Aleut in his Eurasiatic language family. Unlike Illič-Svityč and Bomhard, he does not include Kartvelian, Afrasian, nor Elamo-Dravidian — not because he believes that they are unrelated, but because he believes that these three language phyla are more distantly related to Indo-European than are the others, which, along with Indo-European, form a natural taxonomic subgrouping. Bomhard's opinion is close to that of Greenberg. As he sees the situation, Nostratic includes Afrasian, Kartvelian, and Elamo-Dravidian as well as Eurasiatic; in other words, he views Nostratic as a higher-level taxonomic entity. Afrasian stands apart as an extremely ancient, independent branch — it was the first branch of Nostratic to separate from the rest of the Nostratic speech community. Younger are Kartvelian and Elamo-Dravidian. It is clear from an analysis of their vocabulary, pronominal stems, and morphological systems that Indo-European, Uralic-Yukaghir, Altaic, Gilyak, Chukchi-Kamchatkan, and Eskimo-Aleut are more closely related as a group than any one of them is to Afrasian, Kartvelian, and Elamo-Dravidian, and this is the reason that Bomhard follows Greenberg in setting up a distinct Eurasiatic subgroup within Nostratic. Finally, mention should be made of Sumerian, which Bomhard had investigated in previous works as a possible Nostratic daughter language. Bomhard now believes that Sumerian is not a Nostratic daughter language but that it is related to Nostratic as a whole. It must be noted here that Bomhard has also changed his mind about the subgrouping of Kartvelian and Elamo-Dravidian. His present thinking is that Kartvelian is probably closer to Eurasiatic than what he indicated in his 1994 co-authored

book and that the differences are due to the fact that Kartvelian became separated from Eurasiatic at a very early date. On the other hand, he now sees Elamo-Dravidian as the second group (after Afrasian) to split from the rest of the Nostratic speech community. Bomhard's current views on subgrouping are shown in the following chart:

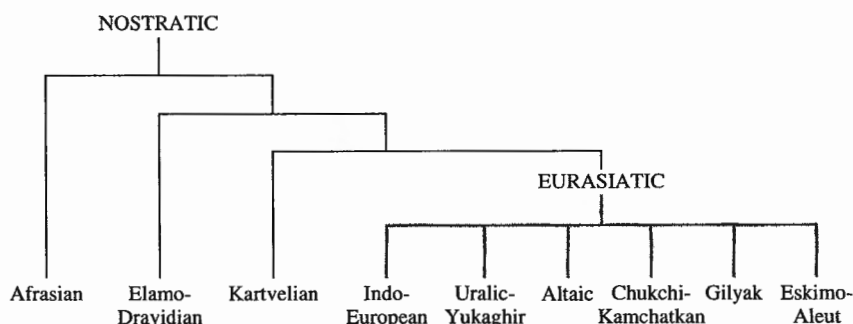


Chart 1. The Nostratic Macrofamily

Interest in issues dealing with Nostratic has resulted in several recent conferences, the first of which was held in Moscow in 1972 to coincide with the publication of the first volume of Illič-Svityč's comparative Nostratic dictionary. This was followed by a series of gatherings in Russia. Another major conference was held in Ann Arbor, Michigan, at the end of 1988. Organized by Vitaly Shevoroshkin and Benjamin Stolz, this symposium brought together scholars from East and West. A series of volumes under the editorship of Shevoroshkin has been appearing as a result of this conference (published by Brockmeyer in Bochum, Germany). Shevoroshkin has also organized several smaller-scale, follow-up conferences. At the end of 1993, a workshop with the theme "The Second Workshop on Comparative Linguistics. The Status of Nostratic: Evidence and Evaluation" was organized at Eastern Michigan University, Ypsilanti, Michigan. Papers from this workshop have just been published in a volume co-edited by Brian Joseph & Joe Salmons (1998). Several important papers on Nostratic also appear in the recent festschrift for Vitalij Shevoroshkin (1997). In December 1997, a workshop on distant linguistic relationship was held at the Santa Fe Institute in Santa Fe, New Mexico, which brought together scholars from around the world.

In early 1998, Dolgopolsky's book entitled *The Nostratic Hypothesis and Linguistic Paleontology* was published. In this book, Dolgopolsky is mainly concerned with linguistic paleontology, and the focus of his attention, therefore,

is on putative etyma pertaining to habitat, social organization, and material culture. Dolgopolsky's conclusions are supported by a sample of 125 proposed cognate sets. The book ends with a reconstruction of the Proto-Nostratic phonological system and the reflexes of the consonants (but not the vowels) in the major branches of Nostratic. This book was the focus of a two-day symposium held in July 1998 at Cambridge University, England.

Author's address

Allan R. Bomhard
151 Wentworth Street, #3B
Charleston, SC 29401-1743, USA
bomhard@aol.com

References

- Bomhard, Allan R. 1984. *Toward Proto-Nostratic: A new approach to the comparison of Proto-Indo-European and Proto-Afroasiatic*. (= *Current Issues in Linguistic Theory*, vol. 27.) Amsterdam: John Benjamins.
- . 1986a. "Common Indo-European/Afroasiatic Roots: Supplement 1". *General Linguistics* 26:4.225–257.
- . 1986b. "The Aspirated Stops of Proto-Indo-European". *Diachronica* 3:1.67–80.
- . 1988. "The Prehistoric Development of the Athematic Verbal Endings in Proto-Indo-European". *A Linguistic Happening in Memory of Benjamin Schwartz* ed. by Yoël L. Arbeitman, 475–488. Louvain-la-Neuve: Publications Linguistiques de Louvain.
- . 1990. "A Survey of the Comparative Phonology of the So-Called 'Nostratic Languages'". *Linguistic Change and Reconstruction Methodology* ed. by Philip Baldi, 331–358. The Hague: Mouton de Gruyter.
- . 1991. "Lexical Parallels between Proto-Indo-European and Other Languages". *Studia Etymologica Indoeuropaea: Memoriae A. J. Van Windekens (1915–1989) Dedicata* ed. by L. Isebaert, 47–106. Leuven: Peeters.
- . 1992a. "Distant Linguistic Comparison and the Nostratic Hypothesis". *FAIES Newsletter* 2(1) (March 1992).
- . 1992b. "The Nostratic Macrofamily (with Special Reference to Indo-European)". *Word* 43:1.61–83.
- . 1992c. "'Zakon stertevanta' v xetsskom: reinterpretacija" [Sturtevant's Law in Hittite: A Reassessment]. *Voprosy Jazykoznanija* 1992:4.5–11.
- . 1993. "Linguistic Methodology and Distant Linguistic Comparison". *Mother Tongue* 20.1–4.

- . 1996. *Indo-European and the Nostratic Hypothesis*. Charleston, South Carolina: Signum Desktop Publishing.
- Bomhard, Allan R. & John C. Kerns. 1994. *The Nostratic Macrofamily: A study in distant linguistic relationship*. Berlin & New York: Mouton de Gruyter.
- Brunner, Linus. 1969. *Die gemeinsamen Wurzeln des semitischen und indogermanisches Wortschatzes* [*The Common Roots of the Semitic and Indo-European Vocabularies*]. Bern: Francke Verlag.
- Collinder, Björn. 1934. *Indo-uralisches Sprachgut: Die Urverwandschaft zwischen der indoeuropäischen und der uralischen (finnisch-ugrisch-samojedischen) Sprachfamilie* [*Common Indo-Uralic Linguistic Possessions: The primitive relationship between the Indo-European and Uralic (Finno-Ugrian-Samoyed) language families*]. Uppsala: A. B. Lundequistska.
- . 1954. "Zur indo-uralische Frage" ["Concerning the Indo-Uralic Question"]. *Acta Universitatis Upsaliensis* 10. 79–91.
- Cuny, Albert. 1946. *Invitation à l'étude comparative des langues indo-européennes et des langues chamito-sémitiques*. Bordeaux: Editions Biere.
- Dolgopolsky, Aaron. 1984. "On Personal Pronouns in the Nostratic Languages". *Linguistica et Philologica. Gedenkschrift für Björn Collinder (1894–1983)* ed. by Otto Gschwantler, Károly Rédei & Hermann Reichert, 65–112. Vienna: Wilhelm Braumüller.
- . 1989. "Problems of Nostratic Comparative Phonology (Preliminary Report)". *Reconstructing Languages and Cultures* ed. by Vitaly Shevoroshkin, 90–98. Bochum: Brockmeyer.
- . 1992. "The Nostratic Vowels in Indo-European". *Nostratic, Dene-Caucasian, Austric and Amerind* ed. by Vitaly Shevoroshkin, 298–331. Bochum: Brockmeyer.
- . 1993. "Nostratic". *The Encyclopedia of Language and Linguistics* ed. by R. E. Asher, volume 5, 2838. Oxford: Pergamon Press.
- . 1998. *The Nostratic Hypothesis and Linguistic Paleontology*. Cambridge: The McDonald Institute for Archaeological Research.
- Gamkrelidze, Thomas V. & Vjačeslav V. Ivanov. 1984. *Indoevropskij jazyk i indoevropcejsky: Rekonstrukcija i istoriko-tipologičeskij analiz prajazyka i protokultury* [*Indo-European and the Indo-Europeans: A reconstruction and historical typological analysis of a protolanguage and a proto-culture*]. 2 vols. Tbilisi: Publishing House of the Tbilisi State University.
- . 1995. *Indo-European and the Indo-Europeans: A reconstruction and historical typological analysis of a protolanguage and a proto-Culture*. 2 vols. English translation by Johanna Nichols. Berlin & New York: Mouton de Gruyter.
- Greenberg, Joseph H. Forthcoming. *Indo-European and its closest relatives: The Eurasiatic language family*. 2 volumes. Stanford, California: Stanford University Press.
- Hegedűs, Irén. 1992. *Bibliographia Nostratica 1960–1990 (A List of Publications on, or Relevant for Nostratic Studies)*. Szombathely: Seminar für uralische Philologie der Berzsényi-Hochschule.

- Hegedűs, Irén, Peter A. Michalove & Alexis Manaster Ramer, eds. 1997. *Indo-European, Nostratic, and Beyond: Festschrift for Vitalij V. Shevoroshkin*. Washington, DC: Institute for the Study of Man.
- Illič-Svityč, V[ladislav] M[arkovič]. 1965. Materialy k sravnitel'nomu slovarju nostratičeskix jazykov (indoevropskij, altajskij, uralskij, dravidskij, kartvelskij, semito-xamitskij) [Materials for a Comparative Dictionary of the Nostratic Languages (Indo-European, Altaic, Uralic, Dravidian, Kartvelian, Hamito-Semitic)]. *Ėtimologija* 1965.321–373.
- . 1971. *Opyt sravnenija nostratičeskix jazykov (semitoxamitskij, kartvelskij, indoevropskij, uralskij, dravidskij, altajskij)* [An Attempt at a Comparison of the Nostratic Languages (Hamito-Semitic, Kartvelian, Indo-European, Uralic, Dravidian, Altaic)]. 3 vols. Moscow: Nauka.
- Joseph, Brian & Joe Salmons, eds. 1998. *Nostratic: Sifting the evidence*. Amsterdam & Philadelphia: John Benjamins.
- Koskinen, Kalevi E. 1980. *Nilal: Über die Urverwandschaft des Hamito-Semitischen, Indogermanischen, Uralischen und Altäischen* [Nilal: Concerning the relationship of Hamito-Semitic, Indo-European, Uralic, and Altaic]. Helsinki: Akateeminen Kirjakauppa.
- Lamb, Sydney M. & E. Douglas Mitchell, eds. 1991. *Sprung from Some Common Source. Investigations into the Prehistory of Languages*. Stanford, California: Stanford University Press.
- Landsberg, Marge. 1986. *Materials for a Bibliography of Translinguistic Studies*. Bloomington, Indiana: Indiana University Linguistics Club.
- Möller, Hermann. 1906. *Semitisch und Indogermanisch* [Semitic and Indo-European]. Reprinted 1978. Hildesheim: Georg Olms.
- . 1911. *Vergleichendes indogermanisch-semitisches Wörterbuch* [A Comparative Indo-European-Semitic Dictionary]. Reprinted 1970. Göttingen: Vandenhoeck & Ruprecht.
- Pedersen, Holger. 1931. *The Discovery of Language: Linguistic science in the 19th Century* English translation by John Webster Spargo. Midland book edition 1962. Bloomington, Indiana: Indiana University Press.
- Ruhlen, Merritt. 1994. *The Origin of Language: Tracing the evolution of the mother tongue*. New York: John Wiley.
- Shevoroshkin, Vitaly, ed. 1989a. *Explorations in Language Macrofamilies*. Bochum: Brockmeyer.
- . 1989b. *Reconstructing Languages and Cultures*. Bochum: Brockmeyer.
- . 1990. *Proto-Languages and Proto-Cultures*. Bochum: Brockmeyer.
- . 1991. *Dene-Sino-Caucasian Languages*. Bochum: Brockmeyer.
- & Thomas L. Markey, eds. 1986. *Typology, Relationship, and Time*. Ann Arbor, Michigan: Karoma Publishers.

CHAPTER 28

Typology and Reconstruction

New trends in comparative historical and diachronic linguistics

Thomas V. Gamkrelidze

Tbilisi, Georgia

It is a special privilege to participate in the Festschrift volume for my colleague and friend E. F. Konrad Koerner, who has contributed so much to furthering new trends in linguistic science by his own writings and through founding new journals, book series, and collective volumes.

The present contribution, based on our joint monograph on the reconstruction of Proto-Indo-European language and culture (Gamkrelidze & Ivanov 1984, 1995), is a small tribute to his outstanding scholarly activities in the field of linguistic science to which he has devoted so much of his life. We very much hope that one day he will also come up with his own assessment of the above-mentioned monograph and its rôle in the development of historical comparative and diachronic linguistics as a whole in the second half of the present century.

The latter half of the 20th C. has been marked in the history of linguistics by an enhanced interest in problems of diachronic studies, this being a return — on a new methodological basis — to the treatment of problems which arose in the classical comparative Indo-European linguistics. This growing concern with problems of diachronic linguistics stems from the general development of linguistic thought over the past decades. Overcoming the Saussurean antinomy between synchronic and diachronic linguistics, it is striving to build a linguistic theory that would have more explanatory power in comparison to purely synchronic theories of taxonomic grammar built strictly on the basis of empirical linguistic data.

The ultimate aim of diachronic linguistics is the reconstruction of early linguistic states with a view to elucidating the ways of the origin and development of historically attested related language systems. Such a reconstruction of

the parent-language state is attained by a comparison of historically attested cognate language systems and by a retrospective movement from one state to another, earlier state, until the language state is reached from which all the historically attested cognate language systems can be deduced under an assumption of a definite set of structural transformations. Such transformations lead us from the original (basic) language system to later language states which are the result of its structural changes.

Structural transformations postulated in this way may be described as vertical or diachronic (in contrast to horizontal transformations generating new constructions within the same system). With the help of these transformations various related language systems are deduced from theoretically postulated basic structural patterns.

By their explanatory power with respect to the language structure, the vertical or diachronic transformations, deducing historically attested forms of a language from certain theoretical constructs considered to be chronologically earlier stages of these forms (their archetypes), are comparable to the horizontal transformations of transformational-generative grammar which deduce the observable surface-structure constructions of a language from theoretically postulated basic constructions constituting the deep structure of the language.

The problems of modern diachronic linguistics are closely linked with the problems and methods of language typology and the linguistics of universals — one of the main branches of modern linguistic science.

One of the basic tasks of present-day linguistics is to establish a structural isomorphism of languages, to study general principles underlying language structures. Typological studies are called upon to bring this isomorphism to light, to identify the invariant characteristics of languages and to establish language universals on various levels of language structure.

Linguistic research has so far established a number of language universals on various levels of language structure, thus enabling a judgement of the invariant characteristics of natural languages.

An application of the principles of language typology and linguistic universals to theoretical premises of historical comparative linguistics and language reconstruction necessitates a reformulation of the aims and tasks of comparative historical linguistics in general, and of linguistic reconstruction in particular.

The principle of typological plausibility, both synchronic and diachronic, of a postulated model for a proto-language provides a new approach to comparative linguistic studies and calls for a revision of traditional views on reconstructed proto-linguistic systems, in particular those on Proto-Indo-European and its

daughter dialects. On the basis of comparative and typological evidence, the traditional and classical threefold system of PIE stops must be given a phonetic reinterpretation whereby the traditional 'plain voiced' stops might be viewed as glottalized; in tabular form:

I	II	III	reinterpreted as	I	II	III
(b)	bh	p		(p')	b/bh	p/ph
d	dh	t		t'	d/dh	t/th
g	gh	k		k'	g/gh	k/kh

The stop series in the new interpretation must be defined as: 1. glottalized; 2. voiced(-aspirate); 3. voiceless(-aspirate), where aspiration is a phonetically relevant, but phonemically redundant, feature.

Such a system of PIE stops, reconstructed on the basis of a comparison of the phonemic systems of historically attested Indo-European languages, taking account of the characteristics of universally valid relations of domination in the phonological system, appears to be — unlike the traditionally reconstructed system — in full agreement with both synchronic and diachronic typological evidence. The suggested system thus appears more probable than the traditionally reconstructed system of Proto-Indo-European stops.

This interpretation of the three series of PIE stops affords a natural phonological explanation of the functional weakness of the labial phoneme /p'/ of the glottalized series I in PIE, which remained unaccounted for in the traditional theory under the assumption of the feature of voice of Series I.

Under such an interpretation a number of restrictions imposed on the structure of the Proto-Indo-European root is given a natural phonetic-typological interpretation. The absence of roots with voiced stops of the *deg-, *ged- type — a fact well known in classical comparative linguistics, but typologically eluding explanation — finds a natural phonetic explanation in the suggested system of PIE stops with the feature of glottalization in Series I. In view of their articulatory-acoustic peculiarities, glottalized stops or ejectives do not tend to combine with each other within a root, a phenomenon that may be illustrated on extensive typological material (cf. the evidence of Amerindian, African, and Caucasian languages with glottalized consonants).

This revision of the consonantism is known in current Indo-European comparative studies as the Indo-European Glottalic Theory, which takes a new look at the Proto-Indo-European linguistic model and its diachronic transformations into the historical Indo-European languages.

On the glottalic analysis these transformations prove to be totally different from those traditionally assumed. The archaic PIE stop inventory proves to be

closer to those of languages traditionally viewed as having undergone a later consonant shift or *Lautverschiebung* (Germanic, Armenian, Hittite), while languages traditionally considered phonologically conservative (especially Old Indian) prove to have undergone complex phonemic transformations in their consonantism.

The traditionally established trajectories of the transformation of the PIE stops into the phonemic units of the individual Indo-European languages change accordingly, acquiring — in the new interpretation of the PIE phonological system — a reverse direction. The basic “phonetic laws” of classical comparative linguistics, such as Grimm’s Law, Grassmann’s Law, Bartholomae’s Law, etc., are being conceptualized anew, acquiring a different meaning in the light of the new interpretation of the Proto-Indo-European system of stops. The typological approach to linguistic reconstruction necessitates a radical reinterpretation and reformulation of all the basic comparative work in the field of Indo-European.

The IE phonological system in such an interpretation appears to be typologically close to the language systems of the historically adjacent areas, Kartvelian and Semitic. This poses new problems as to the interrelationship of these linguistic systems within a common cultural area.

New methods of structural analysis in diachronic linguistics and a typological approach to linguistic reconstruction led to the advancement of the Glottalic Theory, which has been considered, in view of its fundamentally different interpretation of the PIE linguistic system, a new ‘paradigm’ in Indo-European comparative linguistics, comparable in its consequences for the views on the derivation and developments of the individual Indo-European dialects to the Laryngeal Theory.

However, different from the Laryngeal Theory, the Glottalic Theory does not change anything in terms of sound correspondences between historically attested Indo-European languages, but it entails a complete revision of the prehistory of these languages, advancing new IE “laws” and new diachronic transformation rules deriving the historically attested stages of the cognate languages from a common prehistoric stage.

The Indo-European Glottalic Theory has even been viewed by Manfred Mayrhofer, alongside the Palatalgesetz and the Laryngeal Theory, as a final stage in the process of digression in Indo-European comparative studies from the Old Indian pattern as a model for PIE.

Of positive appraisals of the Glottalic Theory, a comment by Winfred P. Lehmann,¹ dating from 1983, may be mentioned here:

Major contributions of the past five decades have modified extensively the views on Proto-Indo-European phonology presented in the standard handbooks

by Brugmann, Hirt and Meillet. These contributions result, on the one hand, from a different approach to the parent language, on the other, from two far-reaching theories, the laryngeal theory and the glottalic theory ... What had seemed one of the most solid achievements of 19th C. linguistics is now modified in every section.

How different it is from the mood which reigned at the beginning of our century, when Antoine Meillet, summing up his views on the situation in comparative IE linguistics could make the following statement in his famous *Introduction* (1908): "En un sens au moins, il semble qu'on soit parvenu à un terme impossible à dépasser".²

Even the modified version of this statement by Emile Benveniste in the posthumous edition of Meillet's *Introduction* (1937: 479-80) does not change anything about the established view: "Même une trouvaille d'espèce inattendue ... n'a pas renouvelé l'idée qu'on se fait de l'indo-européen; le hittite ... n'oblige à rien changer d'essentiel aux doctrines exposées ici; il éclaire nombre de faits, mais il ne transforme pas la théorie générale ...".

The emergence and further development of the Laryngeal Theory, founded on the method of internal reconstruction, and the advent of the Glottalic Theory, based on the principle of synchronic and diachronic typological verification in comparative reconstruction, have brought Indo-European comparative historical studies out of this theoretical stagnation.

The Glottalic Theory is sharing the fate of the Laryngeal Theory also in that it has from the very beginning not been accepted unanimously by specialists in the field, especially of the older generation, this being an additional testimony to the "paradigmatic" character of the new theory of PIE consonantism, with all its consequences for the whole of Indo-European comparative studies.

I would like to recall in this connection the famous "Max Planck principle" (1983: 13).

Eine neue wissenschaftliche Wahrheit pflegt sich nicht in der Weise durchzusetzen, daß ihre Gegner überzeugt werden und sich als belehrt erklären, sondern vielmehr dadurch, daß die Gegner allmählich aussterben und daß die heranwachsende Generation von vornherein mit der Wahrheit vertraut gemacht ist.

Author's address

Thomas V. Gamkrelidze
Oriental Institute
Georgian Academy of Sciences
Tbilisi 380062
GEORGIA

Notes

1. This quote was taken from a pre-print given to the author by Winfred Lehmann. Neither the author nor the co-editors have been able to verify the exact source of the quote.
2. Neither the author nor the co-editors have been able to verify the exact source of the quote.

References

- Gamkrelidze, Thomas V. & Vjacheslav V. Ivanov. 1984. *Indo-European and the Indo-Europeans. A reconstruction and historical typological analysis of a proto-language and a proto-culture* (in Russian), 2 vols. Tbilisi: Tbilisi University Press.
- . 1995. *Indo-European and the Indo-Europeans. A reconstruction and historical analysis of a proto-language and proto-culture* (English version by Johanna Nichols), 2 vols. Berlin & New York: Mouton de Gruyter.
- Meillet, Antoine. 1937. *Introduction à l'étude comparative des langues indo-européennes*. 8th ed. Paris: Hachette.
- Planck, Max. 1983. *Vorträge und Erinnerungen* (5. Aufl. Stuttgart 1949; Nachdruck Darmstadt: Wissenschaftliche Buchgesellschaft).

CHAPTER 29

Typology and Diachrony of the Middle Voice

Helena Kurzová
South-Bohemian University

1. Introduction

In the flectional Indo-European languages, three categories cooperate in expressing fundamental verbo-nominal relations in a sentence: case of nouns, person, and voice of verbs. The case category was from the beginning a basis for distinguishing the main types of languages as defined from the Indo-Europocentric viewpoint. The Indo-European flectional case was opposed to the agglutinative case and to the analytical (prepositional 'cases') and the isolating (word order) expression of 'case' relations. Also, the category of person with personal markers oriented to only one prominent participant, the subject, had already been typologically evaluated in classical typology and in the traditions of Indo-European linguistics. The Indo-European verb was opposed to verbs in many languages that have more than one participant marked by personal markers, i.e., with (direct or also indirect) object agreement in addition to subject agreement (or agent + patient or also benefactive agreement in the later defined ergative structures). Yet the Indo-European middle or mediopassive was considered a rather idiosyncratic category. Its sense as well as its diachronic origin and history were controversial points: which came first, media tantum/deponents or inflectional oppositions? If the latter, which function of the oppositional middle was the starting point?

With the progress of both Indo-European comparative linguistics and general-comparative linguistics (typology), this Indo-European category was relieved of its isolation. Within Indo-European the affinity of middle and perfect and their possible common origin were detected. In typological research it was shown that the 'voice beyond the passive' or the 'middle' voice is typical of

languages oriented to the subject. This provided the hitherto missing typological evaluation of this category and a definition of its position in the grammar. The Indo-European inflectional middle has parallels in similar categories of other languages within the limits of the defined typological constraint.

Are the semantics and function of this category as well as its diachronic transformations now clear enough? Are the old controversies resolved? What are the points of agreement and controversy in modern research? These are questions to which this is a modest contribution.

2. Middle or middle-passive voice: Ancient Indo-European middle and Modern Indo-European reflexive

The defined 'middles' tend toward a very similar sphere of use, involving various 'pseudoreflexive' actions and various physical and psychical, especially emotional, processes and states which can be covered by the term 'inertive'.¹ The inertive function is a typical function of the middle, in addition to the direct/indirect reflexive and reciprocal functions in the true sense² — and to the passive (if developed and if expressed by means of middle form). Unlike reflexives and passives forming regular active vs. middle-passive oppositions, the middles expressing pseudo-reflexive actions or 'inertive' processes and states are regularly media tantum or 'deponents', i.e. there is no active form of the same verb. This feature is typical of the middle category. Individual languages differ in forming 'oppositional middles' other than reflexives and passives.

Perel'muter (1995) defines eight types of regular semantic oppositions between active vs. middle (mediopassive) forms of the same verb in Ancient Greek:

Active vs. middle in the direct and indirect reflexive and reciprocal functions (1–4, cover term: 'reversive')

Active vs. middle in the passive function (5).

Active/causative vs. middle in the function which may be called 'anticausative' (Perel'muter, however, does not use this term), yet has two different semantic variants in Greek:

'inertive' (6), with middle expressing non-intentional processes and states such as χολόομαι 'to be angry' vs. causative χολώω 'to make angry' or ἐγείρομαι intr. 'to wake' vs. ἐγείρω tr. caus. 'to wake';

'actional anticausative' (7) expressing action performed by a person such as πορεύομαι 'to go' πορεύω 'to make go' or περαιόομαι 'to pass over' vs. tr. caus. περαιώω 'to carry over'.

The middle is also actional in the eighth type of opposition with active and middle opposed as 'conversives', i.e., denoting actions performed by different participants in one and the same situation, cf. δανείζω 'to lend' vs. δανείζομαι 'to borrow', γαμέω 'to take as a wife' vs. γαμέομαι 'to give herself in marriage'. This lexically bound type will not be further analyzed here.

What is typical of Greek, along with the much discussed indirect middle, are 'causative' vs. 'anticausative' oppositions of types 6 and 7 above. Only the lack of any oppositional causative member distinguishes *media tantum* such as οἰχομαι 'to go', χόομαι 'to be angry' from the oppositional anticausatives πορεύομαι 'to go' vs. πορεύω 'to make go', χολόομαι 'to be angry' vs. χολόω 'to make angry'.

Media tantum in the true sense, i.e., without active forms, as well as 'lexical middles' differing from actives by their lexical meaning (such as ἀποκρίνομαι 'to give an answer, to reply' to ἀποκρίνω 'separate, choose') are very frequent in Greek. They can be coordinated with the oppositional middles above as pseudoreflexives, inertives, and actional intransitives.

In addition to personal passives, there are impersonal passives, which, however, are formed in Greek from the verbs of saying such as λέγεται, 'it is said, one says' only. By contrast, in Latin (a language with no indirect reflexive and anticausative oppositions), the so-called impersonal passive is regularly formed from the intransitive verbs like *itur* 'one goes', *bene valetur* 'one is well', *recte vivitur* 'one lives rightly', similarly as in other languages with mediopassive -*r*-forms.³

Essentially the same range of functions — reversive, inertive, actional intransitive, passive (of impersonal and backgrounding type) — can be expressed by the reflexive of some modern Indo-European languages, a category corresponding to the ancient Indo-European middle. Not all languages with reflexive verbs cover the whole functional domain, as is also the case in ancient Indo-European languages. We find all 'reversive' functions of the Greek oppositional middle in the reflexives of languages like Czech and Italian, including the indirect reflexive and pseudoreflexive functions. In agreement with Greek active παρασκευάζω vs. direct or indirect middle παρασκευάζομαι, we have Czech act. *připravuji* vs. dir. refl. *připravuji se* and ind. refl. *připravuji si*⁴ 'I prepare, I prepare (med./refl.) myself for, I prepare (med./refl.) something in my interest'. Similarly we have in Italian the so-called 'medio transitivo' *procurarsi* 'to procure for himself' corresponding to active *procurare* 'to procure'. A lot of other transitive verbs form active vs. indirect reflexive oppositions in both Czech and Italian in accordance with the indirect reflexive of Ancient Greek. In addition, intransitive verbs such as 'to stand', or 'to go' may occur in an indirect reflexive form *stojím si*, *jdu si* in Czech, e.g., in sentences like:

Stojím si tu (Jdu si po cestě) a najednou se kolem mne řítí auto.

'stand:1sg.pres ind.refl. here (go:1sg.pres ind. refl. my way) and all at once a car goes racing by'.

Here, the indirect reflexive (beneficiary) component has an emotional value corresponding to the various uses of the so-called *dativus ethicus*.

Note that transitivity of the indirect reflexive middles is weak; the objects are of an incorporating, internal type in typical phrases like 'prepare (ind.refl.middle) one's expedition, one's home-work'. Thus, the events designated by middle verbs in an indirect reflexive function are in principle one-participant events too, as in other middle functions. Two-participant events and highly transitive verbs are included in the functional domain of voice category only when passive.

Like Greek, both Italian and Czech have a large number of *media/reflexiva tantum* or lexical middles/reflexives. In particular, 'inertive' verbs form a large class of reflexive verbs, some expressing processes like 'breaking', etc. with inanimate subjects. Mostly, however, the 'inertives' express mental and emotional processes and states of animate participants, cf. *verba affectuum* like 'being angry', 'enjoying', etc. The term 'inertive' is truly appropriate for the verbs of 'breaking' and other verbs denoting 'spontaneous' events (as Kemmer 1993 puts it). But the verbs of mental and emotional processes are also non-intentional in their prevailing interpretation, although they can be intentional in some contexts such as 'Don't do it or I will be angry'. Thus, the lack of 'control' joins both groups.

Both Italian and Czech can use the reflexive form in the function of the so-called impersonal passive, not only of the type *si dice, říká se* 'it is said, one says' corresponding to Greek λέγεται, but also in true intransitives corresponding to *-r*-mediopassives of Latin and other languages, cf. Czech *jde se*, Italian *si va* with Latin *itur*, Czech *žije se dobře*, Italian *si vive bene* with Latin *bene vivitur*, Italian *si nasce, si muore* (possible also in Czech) with Hittite *akkiškittari*.⁵ The reflexive passive of Slavic and Romance languages is also used with patient in subject position, but with restricted possibilities of foregrounding the definite and personal patients into subject position. This passive is of the backgrounding type, with non-persons and plurals as typical subjects.⁶

3. The middle (middle/passive) as a grammatical category: definition of its invariant value

What is common to all these functions? Can we define a middle/mediopassive and middle/reflexive as a grammatical category with a certain invariant value?

The typological constraint defined in modern typology, i.e., the occurrence of the middle in languages with subject prominence, provides a basis for such a definition. As already stated in the discussion about the passive, the voice category concerns the relationships between the syntactic and semantic structures of the sentence, the latter being described on a more concrete level with the help of the semantic roles like agent, beneficiary, experiencer, etc. and on a more abstract level with the help of two main actant roles differing in the grade of the agentivity: actor (more agentive) and undergoer (more non-agentive). The true common sense of the middle-passive is signaling that the verb is not oriented to the prototypical subject=agent. This definition is able to include the impersonal passive as well.

The identity of subject=agent is violated by an 'additional' patient or beneficiary role (both are 'undergoer' roles) in direct and indirect reflexive functions (in the sense of 'subject = agent + patient', 'subject = agent + beneficiary'). The subject equals 'undergoer' in the passive function and the subject equals actant with neutralized 'actor-undergoer' roles in the 'inertive' function. The 'inertive' subject of the verbs of breaking, etc. and the human 'experiencer' of the *verba affectuum* are to be distinguished on a more concrete level.

A common feature of all these types of middle-passive is defined by some scholars as 'affectedness' of the subject, see e.g., Klaiman (1988). This generalization of the term, originally applied to the accusative/patient participant only ('affiziert'), is not able to include the middles without subject (impersonals) and the middles without an 'undergoer/affectedness' component, e.g., the middles expressing 'translational motion' such as 'to go, to walk'.⁷

Perel'muter gives an account of the above mentioned active vs. medio-passive oppositions of Greek in terms of the correlations between the syntactic subject and semantic actants, the actor and undergoer (in his terminology actor and goal). He defines the following possible correlations covered by voice oppositions: the active denotes the actor, but not the goal; the passive denotes the goal, but not the actor; the reflexive ('reversive') denotes both the actor and the goal simultaneously; the inertive denotes neither the actor nor the undergoer/goal. This covers six of the active vs. middle oppositions mentioned above. In the seventh (and eighth) types both active and middle denote actors. According to Perel'muter no voice difference is expressed through these two oppositions and they are excluded from the voice category. However, the non-agentive actors of the intransitive verbs like *πορεύομαι* 'to go', etc. are subjects ≠ agents too.

Perel'muter is focused on regular active vs. middle/passive oppositions, and hence the most productive grammatical processes. Lexical middles and *media tantum* are excluded from the voice category on the grounds that they are only

lexical phenomena. In my view, signaling the non-prototypicality of the subject is also a grammatical phenomenon in the case of 'lexical middles' (not forming regular oppositions with active, see above) and of *media tantum*.

Agentivity (+action, +control) and non-affectedness are considered as features of the prototypical subjects in the above definitions of the middle-passive. There is, however, another property of the prototypical subject which must be taken into account and which is also disturbed in the middle-passive voice. A prototypical subject is referentially a definite, autonomous entity, being the external initiator/starting point of an action, intentionally oriented to **external** goal/endpoint. Thus, the prototypical subjects are defined by a cluster of features — 'agentive/intentional, external, non/affected' vs. 'non-agentive, internal, affected'. Yet, we must be aware of the fact that the verb and the event designated by the verb is crucial for the conception of the middle; it is the subject person in the verb which is characterized by the above features. Unlike with active verbs, the actants (actors and undergoers) are involved as **internal** components of the verbal process in the middle-passive verbs. This aspect of the marked diathesis is essential in Benveniste's (1950/1966) definition of the middle. The internal organization of the event is also considered as essential in the revised analysis of the middle by Kemmer (1993, 1994) and Croft (1994). Also in the sentences in which the nominal subject is definite and forms an external 'base of predication' (in the so-called 'categorical statements'),⁸ the subject is internalized by using the middle-passive voice.⁹ Note that the external goal/endpoint orientation and its lack is a distinction valid for intransitive verbs too, e.g., the verbs of motion can be directed to the external effect or viewed as an internal process. Therefore, the internal character of the actant encoded by the subject person is a most important aspect in the definition of the middle-passive, which is able to account also for the impersonal passive.

4. Diachrony of the middle

4.1 *Reflexive and primary 'inertive'*

The defined 'middles' tend toward a very similar sphere of use. Does this agreement imply that diachrony too must be the same and that the middles of different languages have the same source? In principle, this is not a necessary assumption. The typological determination of the middle discovered explains the functional agreement. Further, we have enough examples in the linguistic typology of functionally corresponding categories having a different source

(partially responsible for the semantic differences within a functionally defined category).

The reflexive pronoun in the function of direct or indirect object is clearly a source of the reflexive/middle-passive of modern European languages. We have also clear examples of affixes attached to the verbal root or stem used in both true reflexive and middle functions or examples of middle affixes historically related to reflexive affixes.¹⁰ A good example of the development of the middle function on the basis of indirect reflexive is the Georgian 'sataviso-version' as analyzed by Schmidt (1965) and Stempel (1995: 40–42). The Indo-European middle-passive is, however, not marked by means of stem affixes¹¹ or alternations, but by the type of personal endings, by 'differential subject markers' as Haspelmath (1990: 30) puts it.

The derivation of the marked voice from constructions with reflexive pronouns in modern European languages is part of an analytical grammaticalization: from syntax to analytical morphology (with possible further synthesization). The immediate result of this grammaticalization must be active vs. middle/reflexive opposition. The reflexiva tantum of modern Indo-European languages corresponding to media tantum of ancient Indo-European languages are a secondary development.

In the flectional morphology we encounter another type of grammar and grammaticalization. Here the paradigmization¹² of lexically bound derivational processes is a typical development; hence the grammaticalization goes from the lexically bound derivation to inflection. This is also observable in the character of flectional suffixation which, unlike the regular agglutinative suffixation, is both formally (stem alternations, fusions) and semantically bound on lexical stems. Thus, we have every reason to assume that the differential subject person markers had lexically bound application in the differential, non-oppositional verb classes. The person markers of the subjects \neq agents are applied to the verbs expressing actions only, whereas the verbs expressing physical and psychical processes and states do not have the consonantal series of person marker *m*, *s*, *t* in the singular. There is not full agreement among scholars about the historical relationships of the middle to the perfect, but the existence of personal endings without *m*, *s*, *t* markers is an undeniable fact. These endings are directly attested for the perfect: *-a*, *-tha*, *-e* in the singular, and there is a reasonable assumption that the simpler middle endings without consonants attested at least for the 1st and 3rd persons: *-a(i)* and *-o(i)* are primary. The development of the marked middle-passive endings is well understandable just by the development of the marked oppositional voice from the non-oppositional distinction, i.e., the verbal gender distinction 'active vs. inactive' corresponding to the nominal gender

distinction 'animate vs. inanimate'.¹³ The reconstruction of inactive vs. active verb classes for Indo-European is a typological evaluation based on the identification of perfecto-medium as a category opposed to aoristo-present.¹⁴

The latest contributors to the Indo-European middle, Perel'muter (1995) and Stempel (1996), underline the differences between the perfect as a stative category and middle. However, one point cannot be denied, i.e., that both categories display personal endings other than *m*, *s*, *t* endings of the present and aorist system. If the difference in the third person *e* vs. *o* was from the beginning the distinction of stative/perfect vs. process/middle, neither category had *t* of the active present and aorist system.

The formation of middle endings from this starting point can be explained well in connection with its development from the original process category within an inactive verb to the marked diathesis within an integrated verbal system. The perfect which was grammaticalized as aspecto-temporal stem developed stem marking, whereas the simple endings were preserved. For the middle, which developed into the marked oppositional voice category, the personal endings became more marked in accord with the iconicity of markedness. They became more voluminous and more marked as for the phonemic (consonantal) material used.¹⁵ By analogy with the active endings (note that actives and middles regularly stand side by side in context, in some functions also as coordinated verbs), the person markers *m*, *s*, *t* were applied before the inactive vocalic markers (or vocalic-laryngeal — we do not use here the laryngeal notation for simplicity's sake).

The reconstructed active vs. inactive endings cannot serve any other purpose in the Indo-European subject-oriented verb than to mark the subject roles. Whether we classify this distinction as 'voice' or 'prevoice', it concerns the subject prototypicality. However, the differential subject markers of active vs. inactive verbs were not opposed as non-marked vs. marked endings. Thus the subjects \equiv agents were not conceived as non-marked and basic in the verbal coding.

Thus, the view on the diachrony of middle explaining the rise of reflexive/middle as a reduction of elaborated agentive transitive verbs and two-participant events does not seem to be valid for the Indo-European middle. The process of the development of reflexive/middle goes through the agent and patient identity in the true reflexive and related pseudoreflexive functions to the 'inertive' participant (and to the secondary patient and agent non-identity in the passive). In Indo-European the 'inertive' participants (in the broad sense used by Perel'muter, the human experiencer included) were primary denotational correlates of the inactive person markers.

Within one and the same lexical category we can follow the development of oppositional middles. In both examples quoted above and many other cases, the old root verb is medium tantum, whereas the more recent formations are oppositional middles (i.e., forming an opposition with causative/active member), cf. οἶχομαι 'to go' with the denominative verb pair πορεύω vs. πορεύομαι 'to make go' vs. 'to go'. Thus, we encounter the primary 'inertives' and 'motional anticausatives' in Greek such as the verbs above, to which the secondary oppositional causatives were formed in some cases. By contrast, the reflexive verbs of translational motions are more or less lexicalized reflexives/middles of the primary transitive verbs, cf. Czech *vrátit se* 'to go back, to return', Italian *recarsi*, Czech *odebrat se*, 'to process'. Thus, in the Slavic/Romance reflexive/middle we have the 'inertives' and 'motional anticausatives' developed from the true reflexive by means of the neutralization of the actor and undergoer components of their subjects; in Greek we have the primary 'inertives' to which the causative member is formed by secondary development.

4.2 Languages with *-r-passive*: the Latin type of deponent/passive

In some languages an *r*-element was used in forming the marked middle-passive endings. This is good evidence of the non-reflexive source of marked voice. We have every right to consider the *r*-form as a 'deagentivum', indefinite 'man' form: the same element occurs as the 3rd pl. ending in Latin perfect *videre*, *-runt* or OI perfect *vidur*, etc. In Celtic the passive is of impersonal character; in Latin the impersonal and backgrounding function is prominent; impersonal use is attested also for the Hittite middle-passive, etc.¹⁶

In Latin the 'deagentive' 'man' function is prominent in the use of deponent/passive. As in other parts of Latin grammar, the paradigmaticization avoids functional ambiguity, so that the *r*-form of individual verbs has either the function of deponent (distinctive class marker) or the function of passive — personal and impersonal. *Moritur* and *vivitur* are good examples of the verbs able to be used in deagentive sentences by their semantics,¹⁷ but only *vivitur* is so used in Latin. For *moritur* the impersonal use is blocked by the deponential character of the verb. Shall we assume that the paradigmaticization is secondary and reconstruct the impersonal use for *moritur* too? This would be the simplest solution for understanding the process of extending the *r*-marker from the deagentive use to the deponent use. In any case the difference between *moritur* and *vivitur* is additional evidence for the primary media tantum/deponents with the 'inertive' value.

5. Passive: Conclusion and retrospective

The concentration of the latest research on the middle functions should not obscure the fact that it is the passive function which is the most effective grammatical function of the marked voice. It concerns the basic two-participant events and highly transitive agentive verbs, allowing the seeing of these actions from the part of the patient and so changing the perspective of seeing the same events. No wonder that it is also a most widespread voice category across languages.

The advanced grammaticalization of oppositional middle results in passive. The oppositions of non-marked and marked voice of some verbs imply passive interpretation as the unique or most natural interpretation as with the verbs of persecuting, striking, and killing. Another source of the passive lies in the 'inertives' with inanimate subjects such as 'door opens', etc.¹⁸

Both modern research in general comparative linguistics as well as progress in Indo-European studies joined to substantiate Delbrück's view on the diachrony of the Indo-European middle:

Ferner ist wohl einleuchtend, dass der Ausgangspunkt nicht bei den Verben zu suchen ist, welche sowohl aktivische als mediale Formen haben, da die Vermutung nahe liegt, dass der Gegensatz der beiden Genera sich bei ihnen zu grösserer Schärfe zugespitzt habe, sondern bei den aktiva und media tantum. Delbrück (1897: 415).

Denn es scheint, dass nicht selten zu medialen Verben aktive Formen hinzugebildet worden sind, sei es mit derselben Bedeutung, sei es (was das häufigere est) mit kausativ-transitiver. Delbrück (1897: 417).

Wackernagel 1926: 129 also acknowledges the process of forming oppositional middles from original media tantum:

Ausserordentlich vielen Verben, die zuerst nur medial flektiert wurden, ist später ein Aktiv beigelegt worden. Auf diese Weise ist die Zahl der Deponentia reduziert und Doppelflexion eingeführt.

The often mentioned controversy between Delbrück and Wackernagel concerns the different approaches to the definition of the functional value of the middle. The good starting point ('Ausgangspunkt' above) is in Delbrück's view the media tantum and in Wackernagel's view the oppositions. The diachronic process of forming regular oppositions from an irregular lexically bound formative base was assumed by both scholars. They did not, however, have the necessary conceptual background for the radicalization and full understanding of this process. This is provided by the cooperation of modern typology and historico-

comparative research, whose results are the definition of active and inactive verb classes and the reconstruction of active, i.e., presento-aoristic and inactive, i.e., perfecto-medial personal endings, for Proto-Indo-European.

With the concepts and tools of modern Indo-European research we can better understand the processes presupposed by Delbrück. The active vs. middle oppositions were formed primarily in inactive verbs whose originally ambiguous or vague meaning was thus disambiguated: the diathetically vague 'burning' splits into the transitive action of 'burning' and the intransitive middle (and intr. perfect); cf. Delbrück (1897: 36–37, 418) for the verb *δαίω, δαίωμαι*. It is the analogy with transitive active/agentive verbs which gave rise to these new transitive active presents, and as a counterpart of this process, the middle-passives are formed from active/agentive verbs like *φέρω, διώκω*, etc. The important component of this reorganization of the Indo-European verb is the thematic present as a common category of inactive and active verbs:¹⁹ 1st sg. *-ō* can be explained as a thematic vowel *o* followed by the inactive marker *H* (*h₂*), for 2nd and 3rd sg. the inactive 2nd/3rd sg. *-e-i* (Greek and Lithuanian evidence supported by Slavic and Celtic facts) and the active endings *-es-i* and *-et-i* were probably allomorphic variants in some period of Indo-European linguistic history. These and similar special solutions will remain points of disagreement among scholars due to the ambiguous character of the comparative evidence for reconstructing the non-attested stages of Indo-European development. This, however, does not devalue the progress in understanding the sense and diachrony of the middle.

Acknowledgments

This research was supported by the Grant Agency of the Academy of Sciences of the Czech Republic, grant No. AO 142601.

Author's address

Helena Kurzová
Václavkova 14
160 00 Praha 6
THE CZECH REPUBLIC

Notes

1. This is the term used by Perel'muter 1995.
2. I.e., in cases where non-reflexive objects also occur in the same meaning.
3. For impersonal uses of the Hittite mediopassive see the examples by Neu 1968b; e.g., the impersonal *akkiškinari* is used about 'mass dying': see Neu (1968b:2), Ivanov (1963:163). The impersonal use of the mediopassive is not restricted to the languages with *-r*-passive; it is widely used in Old Indian. Its limited use is characteristic of Greek.
4. The direct reflexive and indirect reflexive element have distinct forms in Slavic.
5. See note 3 above. Concerning Latin *moritur, nascitur*, see below.
6. This type of passive is described in many contributions in Slavic and Romance linguistics. See also Kurzová (1990).
7. See examples by Kemmer (1993:18), and the Greek examples above.
8. See Sasse (1987).
9. See Croft (1994) for the 'self contained' character of each event encoded by the single verb.
10. See examples in Kemmer (1993, 1994).
11. Stem suffixes developed into the markers of diathesis and the related (valency changing) categories in individual Indo-European languages.
12. The process of forming inflectional paradigms from lexically bound derivations; see Kurzová 1993 and recently Hoskovec 1996.
13. Note that the existence of these verbal classes does not imply the active syntactic structure in Indo-European which is a language structure with unipersonal verb; the nominative-accusative syntax is a primary Indo-European structure. For more see Kurzová 1993.
14. For perfect-medium see especially Neu (1968a, 1976), Meid (1975, 1979); for inactive vs. active classes Schmidt (1979, 1986), Gamkrelidze & Ivanov (1984/1995), Lehmann (1995).
15. Two components of formal markedness are magnitude (voluminosity) and heaviness for the phonemic material used.
16. The middle-passive ending of *-tor* or *-or* type (without the secondary analogic *t*-element marker) further occurs in Osco-Umbrian, Tocharian, Venetic, Armenian, and Phrygian. On *-r*-endings cf. especially K.H. Schmidt (1963a:167), Neu (1968:8ff., 161ff.), Jankuhn (1969:30ff.), G. Schmidt (1977), and Solta (1984 [1985]:65–66).
17. See the above examples for Italian and the Hittite form *akkiškitari*.
18. The development of the passive from middle is well described by Jankuhn (1969).
19. The central position of this present category in the verbal system is a non-deniable fact, but the explanation of thematic present varies; cf. Kurzová 1993.

References

- Benveniste, Émile. 1950. "Actif et moyen dans le verbe". *Journal de Psychologie* 43.121–129 (= Benveniste 1966:168–175).
- . 1966. *Problèmes de linguistique générale*. Paris: Gallimard.

- Croft, William. 1994. "Voice: Beyond control and affectedness". Fox & Hopper 1994.89–117.
- Delbrück, Berthold. 1897. *Vergleichende Syntax der indogermanischen Sprachen*. Vol.II. Strassburg: Karl Trübner.
- Fox, Barbara & Paul J. Hopper, eds. 1994. *Voice: Form and function*. (= *Typological Studies in Language*, 27.) Amsterdam & Philadelphia: John Benjamins.
- Gamkrelidze, Thomas V. & Vjačeslav V. Ivanov. 1984. Индоевропейский язык и индоевропейцы: Реконструкция и историко-типологический анализ праязыка и протокультуры. Vol. I. Tbilisi: Izdatel'stvo tbilisskogo universiteta.
- . 1995. *Indo-European and the Indo-Europeans. A reconstruction and historical analysis of a proto-language and a proto-culture*. Part I The Text, transl. by Johanna Nichols. (= *Trends in Linguistics. Studies and Monographs*, 80.) Berlin & New York: Mouton de Gruyter.
- Haspelmath, Martin. 1990. "The Grammaticalization of Passive Morphology". *Studies in Language* 14:1.25–72.
- Hoskovec, Tomáš. 1996. "Le temps comme catégorie verbale et la reconstruction de l'indo-européen". *Baltistica* 31:1.9–22.
- Ivanov, Vjačeslav V. 1963. Хеттский язык. Moskva: Izdatel'stvo vostočnoj literatury.
- Jankuhn, Harald. 1969. *Die passive Bedeutung medialer Formen untersucht an der Sprache Homers*. Göttingen: Vandenhoeck.
- Kemmer, Suzanne. 1993. *The Middle Voice*. (= *Typological Studies in Language*, 23.) Amsterdam & Philadelphia: John Benjamins.
- . 1994. "Middle Voice, Transitivity, and the Elaboration of Events". Fox & Hopper 1994.179–230.
- Klaiman, M. H. 1988. "Affectedness and Control: A typology of voice systems". *Passive and Voice* ed. by Masayoshi Shibatani. (= *Typological Studies in Language*, 16.), 25–83. Amsterdam & Philadelphia: John Benjamins.
- Kuržová, Helena. 1990. "Die Diathese in den indogermanischen Sprachen". *Croatica Slavica. Indoeuropaea* ed. by Georg Holzer. (= *Wiener Slavistisches Jahrbuch. Ergänzungsband* 8.), 121–130. Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- . 1993. *From Indo-European to Latin. The evolution of a morphosyntactic type*. (= *Current Issues in Linguistic Theory*, 104.) Amsterdam & Philadelphia: John Benjamins.
- Lehmann, Winfred P. 1995. *Residues of Pre-Indo-European Active Structure and their Implication for the Relationships among the Dialects*. (= *Innsbrucker Beiträge zur Sprachwissenschaft, Vorträge und Kleine Schriften*, 61.). Innsbruck: Institut für Sprachwissenschaft der Universität.
- Meid, Wolfgang. 1975. "Probleme der räumlichen und zeitlichen Gliederung des Indogermanischen". *Flexion und Wortbildung: Akten der 5. Fachtagung der Indogermanischen Gesellschaft, Regensburg 9.-14. September 1973* ed. by Helmut Rix, 204–221. Wiesbaden: Ludwig Reichert.

- . 1979. "Der Archaismus des Hethitischen". *Hethitisch und Indogermanisch. Vergleichende Studien zur historischen Grammatik und zur dialekt-geographischen Stellung der indogermanischen Sprachgruppe Altkleinasiens* ed. by Erich Neu & Wolfgang Meid, 159–176. Innsbruck: Institut für Sprachwissenschaft der Universität.
- Neu, Erich. 1968a. *Das hethitische Mediopassiv und seine indogermanischen Grundlagen*. Wiesbaden: Otto Harrassowitz.
- . 1968b. *Interpretation der hethitischen mediopassiven Verbalformen*. Wiesbaden: Otto Harrassowitz.
- . 1976. "Zur Rekonstruktion des indogermanischen Verbalsystems". *Studies in Greek, Italic, and Indo-European Linguistics offered to L. R. Palmer* ed. by Anna Morpurgo Davies & Wolfgang Meid, 239–254. Innsbruck: Institut für Sprachwissenschaft der Universität.
- . 1985. "Das frühindogermanische Diathesensystem: Funktion und Geschichte". *Grammatische Kategorien. Funktion und Geschichte. Akten der VII. Fachtagung der Indogermanischen Gesellschaft, Berlin, 20.- 25. Februar 1983* ed. by Bernfried Schlerath & Veronica Rittner, 275–295. Wiesbaden: Ludwig Reichert.
- Perel'muter, Ilja Aronovič. 1995. Залог древнегреческого глагола: Теория, генезис, история (Engl. sum. 248–262: Voice in the Ancient Greek Verbal System: Theory, Genesis, History). Saint Petersburg: Izdatel'stvo "Noach".
- Sasse, Hans-Jürgen. 1987. "The Thetic/Categorical Distinction Revisited". *Linguistics* 25.511–580.
- Schmidt, Gernot. 1977. "Das Medium im vorhistorischen Keltisch". *Indogermanisch und Keltisch. Kolloquium der Indogermanischen Gesellschaft am 16. und 17. Februar 1976 in Bonn* ed. by K.H. Schmidt & Rolf Ködderitzsch, 89–107. Wiesbaden: Ludwig Reichert.
- Schmidt, Karl Horst. 1963a. "Venetische Medialformen". *IF* 68.160–169.
- . 1963b. "Zum altirischen Passiv". *IF* 68.257–275.
- . 1965. "Indogermanisches Medium und Sataviso im Georgischen". *BK* 19–20 (48–49): 129–135.
- . 1979. "Reconstructing Active and Ergative Stages of Pre-Indo-European". *Ergativity. Towards a theory of grammatical relations* ed. by Frans Plank, 333–345. London: Academic Press.
- . 1980. "Zur Typologie des Vorindogermanischen". *Linguistic Reconstruction and Indo-European Syntax. Proceedings of the Colloquium of the "Indogermanische Gesellschaft", University of Pavia, 6–7 September 1979* ed. by Paolo Ramat, 91–112. Amsterdam: John Benjamins.
- . 1986. "Zur Vorgeschichte des prädikativen Syntagmas im Indogermanischen". *O-o-pe-ro-si. Festschrift für Ernst Risch* ed. by Annemarie Etter, 90–103. Berlin: Walter de Gruyter.
- Solta, Georg Renatus. 1984[1985]. Review of Gert Klingenschmitt, *Das altarmenische Verbum*. (Wiesbaden: Ludwig Reichert, 1982). *Kratylos* 29.59–74.

- Stempel, Reinhard. 1996. *Die Diathese im Indogermanischen. Formen und Funktionen des Mediums und ihre sprachhistorischen Grundlagen.* (= *Innsbrucker Beiträge zur Sprachwissenschaft, Vorträge und Kleinere Schriften*, 67). Innsbruck: Institut für Sprachwissenschaft der Universität.
- Wackernagel, Jacob. 1926. *Vorlesungen über Syntax mit besonderer Berücksichtigung von Griechisch, Lateinisch und Deutsch.* 1. Reihe. 2nd ed. Basel: Birkhäuser. (First ed., 1920.)

CHAPTER 30

Indo-European Numerals Since Szemerényi

Carol F. Justus

University of Texas at Austin

1. Introduction

Konrad Koerner's contributions to the history of Indo-European (IE) linguistic traditions are well known. Szemerényi (1960), following 19th C. linguistic traditions, assumed that shared words for digits 1–9, 10, and 100 implied a system with fully developed decades and devised sound correspondences to support the view that the Proto-Indo-Europeans had a decimal system. Although sound correspondences indeed identify older and newer numeral forms as in some sense the same, the meanings of those forms need not therefore be static. Szemerényi's assumptions in fact led him to overlook the possibility that old forms may have changed meaning, that the system which gave meaning to numeral words in earlier millennia might have evolved along with technological innovations.

Recent studies of Ancient Near Eastern accounting show that the conceptual systems underlying counting evolved rapidly at the end of the fourth and beginning of the third millennium BC, the period when writing evolved. If we agree with Gamkrelidze & Ivanov (1983), Mallory (1989: 222–265), and others in dating the Proto-Indo-European homeland to a period before the late fourth millennium BC, then the Homeland predates the onset of both literacy and numeracy. Szemerényi's Proto-IE decimal system is thus at odds with attested evidence for the evolution of numerals and current Homeland dating.

Re-examination of Roman numeral notation beside subtractive strategies for forming the Greek and Latin teen numerals and the Germanic upper decades points to at least three relic bits of data that reflect concepts more akin to those underlying third millennium BC systems of (ac)counting than to the 20th C.

metric system. Early systems of quantification sequenced sizes of collections and subdivided the intervals into unsequenced units. While we now count in terms of sequenced digits, early systems counted in terms of sequenced collection units closer to units such as the foot, yard, and mile, hours of the day, or the units V, X, L, C, and M of Roman numeral notation. Collection units were the sequenced units of counting before digits came to be sequenced.

This study, based on 20th C. contributions to our understanding of early numeral systems, revises Szemerényi's view that the Proto-Indo-Europeans left the Homeland with a decimal system.

2. Szemerényi's thesis

Szemerényi's reconstruction reflects a way of thinking about numerals so intuitive that it is difficult to entertain any other. Devising sound correspondences to derive *k'mtóm "100", *dék'm "10", and decade formative *k'ont / *k'nt all from PIE *dék'mt "ten", earlier *de-kont "two hands", he was able to maintain the view that the decades were largely inherited:

Szemerényi's Origin of '10', '100', and the Decade Formatives

*dék'mt "10" (preconsonantal variant *dék'm*) < *de-k'ont- 'two hands'.

*dék'mt > collocation *dék'mt-dékm'tóm 'ten of tens' > compound

*(d)k'mmk'mtóm '100'.

Antepenult shortening > (d)k'mk'mtóm; haplology > *dk'mtóm;
consonant cluster reduction > *k'mtóm '100' and ablaut variants,
*-k'ont / *-k'nt.

(Szemerényi 1960: 67–69; 139–140).

Given shared IE words for digits 1–9, base 10, and its power 100, Szemerényi concluded that their inner logic must be decimal; therefore, he argued, there must have been fully formed words for the decades 20–90. Even though he acknowledged that there is no common IE form for "1000" (Szemerényi 1960: 2), he assumed that, if the well-known difficulties in reconstructing common decade forms could be solved, the PIE system would be shown to be decimal:

First, it is undeniable that even inherited decad-formations have been replaced by more modern expressions in several IE languages. Thus no one would dare to infer from the fact that the Germanic or Balto-Slavic languages express "20" by juxtaposition of "two tens" that these languages did not inherit a word corresponding to Lat. *viigintii*. Secondly, the fact that the Indo-Europeans already had the number "100" — based on the mathematically well-defined expression "ten tens", and not a vague notion of "indefinite large amount" —

makes it quite clear that there must have been corresponding expressions for "60-90". Sommer is therefore right in refusing to accept this "defectiveness" of the IE system of decads (Szemerényi 1960: 2).

Correspondences among digits 1-9, 10, 100 in fact are striking:

Some IE Digits, Ten, and Hundred Words¹

PIE '1'	Old Irish	Latin	Gothic	English	Lithuanian
*oy-no-	oen	unus	ains	one	vienas
*sem-, *sm-yH-	Greek εἰς, μία	Mycenaean e-me	Armenian mi	Tocharian A sas	
*oy-ko-, *oy-wo	Sanskrit éka	Mitanni aika-(wartanna)	Avestan aeeva-	Old Persian aiva-	

PIE	Sanskrit	Greek	Latin	Old Irish	Gothic	English	Russian
2*dwo(u) [du.]	dwá	δύο	duo	doou	twai	two	dva
3 *tréyes [pl.]	tráyas	τρεῖς	trees	tri	thrija (n.)	three	trie
4*k ^w e-tw(o)r	catvará-	τέτταρες	quattuor	cethair	fidwor	four	chetyre
5*pénk ^w e	pāñca	πέντε	quinque	cooic	fimf	five	pjat'
6*s(w)eks	sás	ἕξ	sex	see	saihs	six	shest'
7 *septm	saptá	ἑπτὰ	septem	secht n-	sibun	seven	sem'
8*októ(u)	astá(u)	ὀκτώ	octo	ocht n-	ahtau	eight	osm'
9 *(e)néwn	náva	ἐννέα	novem	nooi n-	niun	nine	devjat'
10 dék'm	dása	δέκα	decem	deich n-	taihun	ten	desjat'
100 *kmtóm	satám	(ἐ)κατόν	centum	ceet n-	hunda	hundred	sto

The inner logic of a system with these forms and an attested value of "100" as "ten tens", he hypothesized, must be decimal; therefore, there must have been words for fully formed decades 20-90. He says further:

If the IE system of decads is, as can be expected in view of the first members, a logical system, in other words, if it is based on the clear mathematical notions two tens, three tens, etc., then the second member can be predicted to be the numeral "10", even though, perhaps, in a modified form. This *a priori* deduction is borne out by the observable facts. The morpheme *-kont/knt-*, contained in the IE decads "20-90", is obviously related to "10" which itself appears as **dek*m or **dekmt*. We must therefore conclude that *-kont/knt-* derive from earlier **(d)komt/(d)kmt* (Szemerényi 1960: 67).

This conclusion is supported by agreements in form for Old Irish, Latin, Greek,

and Armenian collocations of digit plus *k'ont/*k'nt reflexes (Szemerényi 1960: 5–26):

Decades formed with reflexes of *k'ont/*k'nt

	OIrish	Latin	Greek	Armenian	Gothic	Sanskrit
20	<i>fi-che</i>	<i>vii-gint-i</i>	<i>ἑ-κοσι</i>	<i>k'-san</i>	<i>(stega)</i>	<i>vim-sat</i>
30	<i>tri-cho</i>	<i>trii-gintaa</i>	<i>τριακό-κοντα</i>	<i>ere-sun</i>	<i>(thrije tigiwe)</i>	<i>trim-sat</i>
70	<i>sechtmo-go</i>	<i>septuua-gintaa</i>	<i>ἑβδομή-κοντα</i>	<i>ewt'ana-sun</i>	<i>sibunte-hund</i>	<i>sapta-tí</i>

Szemerényi's (1960: 5–22) phonological solution for anomalies of the decade juncture vowel here is less important than his analysis of the decade suffix *-k'ont- and its ablaut variant *-k'nt- as “ten” forms (1960: 22–26 *passim*). This suffix underlies the Old Irish, Latin, and Greek forms. Armenian too reflects the ablaut patterns with regular palatalization of the original occlusive in prototype forms *-san* and *-sun*. Sanskrit lower decade *-sat* and Germanic upper decade forms in *hund* similarly reflect *-k'nt-. Since forms match down to ablaut variants, Szemerényi (1960: 67) assumed a decimal logic and expected digit plus 10 collocations in the decades. If *k'ont was a 10 form, it fit. But Balto-Slavic had *dék'mt reflexes, not reflexes of *k'ont:

Decades formed with reflexes of *dék'mt

	Lithuanian	Latvian	Bulgarian	Russian
20	<i>dvì-deshimti</i>	<i>div-desmit</i>	<i>dvá-deset</i>	<i>dvá-dcat'</i>
30	<i>trys deshimtys</i>	<i>trīs-desmit</i>	<i>trī-deset</i>	<i>trī-dcat'</i>
70	<i>septynios deshimtys</i>	<i>sept, n-desmit</i>	<i>sedem-desét</i>	<i>sém' desjat</i>
10	<i>deshimt(is)</i>	<i>desmit</i>	<i>deset</i>	<i>désjat</i>
100	<i>shimtas</i>	<i>simts</i>	<i>sto</i>	<i>sto</i>

Szemerényi (1960: 63) viewed these forms, which, like those of Old Icelandic (see below), are not reflexes of *k'ont- / *k'nt-, as “new, clearer types based on 10”. And in Russian *devjanó-sto* “90” where *sto* is a reflex of the old formative *k'ont, the *sto* formant, cognate with Russian *sto* “100” was for Szemerényi (1960: 64–66) archaic. But it was an old “10” reflex of *k'ont where both *k'ont and *dék'mt- were derived from an original PIE *dék'mt “ten”. He rejected earlier analyses of *devjanó-sto* as “ten (wanted) to (make) a hundred” (Endzelin) or “nine to, short of, hundred” (Jagic), saying that it was “difficult to see why the speakers should have thought of ‘9 towards 100’ as a suitable expression for ‘90’” (1960: 64). Just such a subtractive strategy is what we also find in the Greek and Latin teen formations (see below).

For Szemerényi, who reconstructed *dék'mt as original with *k'ont- / *k'nt-

and *k'mtóm all derived from the 10 word, formation of decades on a 10 word would be inherited. The fact that these forms represented newer and older stages had no special significance for him. He could thus ignore the possibility that replacement of older *sto* with newer *desjat* may have reflected changes in numeral systems. He could ignore that in Germanic, Baltic, and Slavic a subdivision of a collection unit might be subtracted from the next upper unit type or that counting might proceed to the next larger unit then double back (e.g., "quarter till five" instead of "four forty-five": "over counting"). In fact such well-known peculiarities have a natural explanation as relics of an old system in which collection units were the sequenced units.

3. Sequencing in emerging systems of counting

Our earliest systems of counting are attested in the clay artifacts called tokens used in the Neolithic across the Ancient Near East as counters for keeping track of quantities of goods. As these clay token shapes became impressed onto clay tablets for record keeping (Schmandt-Besserat 1992; 1996), the conceptual system behind them evolved into the protoliterate systems of numeration. Commodity-specific factor intervals characterized these systems for quantifying different types of commodities (Friberg 1979; Damerow & Englund 1987, 1989; Nissen, Damerow & Englund 1990). Such systems were separated by highly irregular factor intervals, not yet the regular intervals of exponential 10 nor any other exponential factor. Even with more regular sexagesimal numerals, factor intervals alternated between repeated intervals of 6 and 10. Only with regular factor intervals such as decimal 10 do sequenced units become powers.

Mathematicians, distinguishing between concrete and abstract counting, view the properties of cardinality (the abstraction of a notion of 'twoness' or 'threeness' from various collections of two's and three's) and sequence as fundamental to numerical meaning (Danzig 1959). Irregularly sequenced commodity-linked units of the protoliterate systems lacked the mathematical property of cardinality; they did not represent abstract numeral values.

3.1 *Sequenced units*

The decimal system itself might be viewed as sequences of units defined by regular factor intervals of 10:

Views of The Decimal System

Factors:	10	10	10
	_____ ten	_____ hundred	_____ thousand
Powers:	10^1	10^2	10^3

Units separated by such regular intervals as 10, 10, 10 are now interpreted as powers in an exponential system based on 10. The later interpretation of repeated factors 10, 10, 10 would have reanalyzed erstwhile factor intervals between arbitrary collection units named 'ten' and 'hundred' as powers of a base 10. The exponential powers, 10^1 , 10^2 , and 10^3 , are the modern counterparts of units in an earlier system whose values were defined by irregular factor intervals. Friberg (1978; 1979) used these factor diagrams to visualize the irregular intervals that characterized mathematical relations in protoliterate token impressions. Decimal powers undoubtedly standardized the earlier irregular factor intervals that once separated arbitrary units named 'ten', 'hundred', and 'thousand'. Between units defined by irregular factor intervals and powers lie conceptual developments fundamental to changes in the meanings of forms associated with 'ten' and 'hundred'.

Irregular intervals between early collection units, not yet the regular intervals of a repeated 10, thus determined the meaning of early linguistic units. If PIE numerals developed as early as the early third millennium BC, linguistic forms corresponding to 'ten' and 'hundred', if mathematical, would have named units sequenced by irregular factor intervals, not accumulations of sequenced digits spaced by powers of 10. When the system changed, the old unit names were sometimes retained (Old Irish, Greek, Latin, Armenian, Germanic), sometimes recreated (Slavic). Forcing both *k'ont and *dek'mt- reflexes phonologically back to a single 'ten' form as Szemerényi did obscures reinterpretations that the IE numerals underwent.

To our 20th C. mind irregularly sequenced units are more like traditional measure units than like digits or powers of a base. Units 'foot' and 'yard' in length measure and 'cup', 'pint', 'quart' capacities are defined by irregular factors 12 and 3 or 4, 2, 2:

Traditional Measure Units

Length	3	12	
	yard _____	foot _____	inch
Capacity	4	2	2
	gallon _____	quart _____	pint _____ cup

We call these 'measure units', but in evolutionary terms, they share many properties with the pre-numerical units of the protoliterate systems of numeration and the token shapes used even earlier for record keeping. No one would call measure units bases or powers, nor confuse them with abstract numerals. Distinct from abstract numerals in terms of the kind of unit sequenced, they nevertheless count individual objects, but the objects are in collections of smaller units. Being sequenced, but lacking abstract numeral values, the foot and quart lie between the one-to-one correspondence of the mathematician's concrete counting and the exponential decimal numerals we use for mathematical calculations.

Linguistic study of counting sequences, assuming an orientation primarily in terms of digit sequences from one to infinity, focuses on mathematically regular processes of addition and multiplication as characteristic strategies for building higher numerals. Hurford's (1987: 247–250) term 'higher numerals' lumps teens (formed by addition), decades (formed by multiplication), and powers of a base together as a class. Yet all reflect layers of numeral development. Powers of base 10 ('ten, hundred, thousand, million') are a recent type of sequenced unit. Syntactic 'packing' strategies (Hurford 1987: 242–261) account for linguistic terms for teens, decades, hundreds, thousands which presuppose processes of addition and multiplication. Yet addition and multiplication reflect different kinds of unit sequencing. Multiplication sequences collections of units, decades, hundreds, thousands (e.g., multiplicative 'for-ty, four hundred'), while addition sequences digits. The 'packing strategy' for teen number words, for example, reflects sequences of digits from '11' to '19', while the strategy for packing units of decades, hundreds, and thousands 'packs' sequences of units, 10, 100, and 1000. The analysis here suggests that decade conceptualization at a stage with pre-numerical values more akin to early tokens and protoliterate units of numeration viewed them, not as multiplication, but interval measurements between a lower and a higher size of sequenced unit. Decades subdivided (i.e., 'subitized': cf. Høyrup 1994) the interval between 'ten' and 'hundred'. Similarly teens subdivided the interval between one 'ten' unit and the next. Teen digits and decade units were not initially based on addition or multiplication, but on division.

Hurford's synchronic focus gives rules that will generate attested linguistic numeral phrases without accounting for the strategies attested in the evolution of counting from token accounting through protoliterate systems of numeration and Sumerian sexagesimal numerals to the decimal system. Stages in the evolution of quantification include:

unit in a system sequenced by factor intervals 10 and 12. Ulf-Møller (1991 *passim*) has shown that calculations underlying texts such as the Anglo-Saxon Chronicle are based on a long hundred value 120, not the decimal power of base 10. One passage, for example, refers to three-hundred-five days in a year, i.e., three long hundreds ($120 \times 3 =$ decimal 360) plus 5, or decimal 365:

A Conversion between 'Hundred's'

LONG HUNDRED DECIMAL HUNDRED

305 days in a year 365 days in a year

hund of 120 *hund* of 100

$x^3 = 360$ $x^3 = 300$

+ 5 digits = 365 + 65 digits = 365

In terms of evolutionary stages, the difference between the decimal hundred and the long hundred is not a matter of bases but of sequencing strategy: the *hund* of the long hundred is sequenced as second in a system with interval factors 10 and 12. Sommer (1951) grappled unsuccessfully with this distinction. Following Schmidt (1890: 24; 40–43) he realized that an exponential duodecimal system based on 12 (12 times 12) would have yielded a hundred of 144:

Duodecimal Exponentiation

Factors:	12	·	12	
	_____ twelve _____			
Powers:	12^1		12^2	hundred

The fact that the Germanic long hundred of 120 was not duodecimal and that one of its factors was 10 convinced him that it must be decimal. Rejecting Schmidt's suggestion of influence from the Sumero-Babylonian system based on powers of 60, he did not conceive that it might be neither decimal (based on exponential powers of 10) nor duodecimal (based on an exponential 12), but a set of collection units sequenced by factors, 10 and 12. Such a system may have been too reminiscent of measurement systems for Sommer to entertain that early Germanic counting could have used pre-exponential calculation techniques.² When he concluded that the long hundred was decimal, however, he had no way to talk about the conceptual differences in sequencing that separate collection units and powers, hence his confusion. Szemerényi, following Sommer, also assumed counting to be based on exponentially sequenced powers. Since digits added up to 10 linguistically (example 2 above) and decades were sequences of 10, he devised a linguistic analysis that would account for 'common sense' decimal thinking.

Similarly, accounting synchronically for the linguistics of decades and other higher number word forms, based as they are on multiplicative and additive

strategies, undoubtedly steered Hurford (1987: 240–261) away from potential diachronic differences in the basis of sequence (e.g., digits as opposed to collection units). Decades, hundreds, thousands are after all powers of 10. Who would think that they were once the sequenced units subdivided by unsequenced units?

3.3 *Other systems of collection units*

Numeral systems with sequenced collection units instead of powers of a base are in fact numerous and may make use of widely varying factor intervals. Whereas Germanic sequenced collection units at factor intervals 10 and 12, the Sumerio-Babylonian sexagesimal units *u*, *ñésh*, *ñésh-u*, and *s(h)ár* were separated by intervals 10 and 6:

Sumero-Babylonian Sexagesimal Factors

Factors:	6	10	6	10	
	<i>s(h)ár</i>	<i>ñésh-u</i>	<i>ñésh</i>	<i>u</i>	<i>dish</i>
Numerals:	3600	600	60	10	1

Ten of linguistic form *dish* makes the first collection unit *u*, and 6 of *u* results in the second collection unit, *ñésh*, and so on. Sumerian sexagesimals operated with repeated factors 10 and 6, so that units *u*, *ñésh*, *ñésh-u*, and *shár* are not powers of either 10 or 6. At the same time, unlike the foot and yard, they have abstract numeral values. Also unlike the foot and quart, *ñésh* ‘60’ is not limited in its use to any one kind of commodity, so that it has numerical value. It is a ‘collection unit’, not a power of a base.

Schmidt’s (1890) hypothesis, that the break between 60 and 70 in the Germanic long hundreds and elsewhere resulted from contact with the Mesopotamian sexagesimal system, needs further examination now. Both Germanic and pre-Babylonian systems are typologically similar in that both are built on collection units sequenced by factor intervals, not exponentially on powers of a base. After this typological similarity, however, differences also obtain, as Schmidt (1890: 43–53) was aware. Although the sexagesimal system is now built on powers of 60, in its origins it was not (Powell 1997), but on intervals that recalled its more irregular history.

Earlier protoliterate systems of numeration, like the token system of accounting that they followed, were sequenced by highly irregular factor intervals. Powell (1989; 1995) has detailed uses of different factor intervals for calculating seed grain, yield for food, and distribution of rations, often different in cities that traded with each other, and has shown that an apparent chaos of quantification systems led to standardization of old commodity-linked factor

intervals in the Sumerian sexagesimal system. To solve practical mathematical problems such as the amount of seed needed to sow a field in order to produce enough grain to feed a particular population size, standardization of commodity-linked sets of factors eliminated cumbersome conversions from one system of factor sets to another (one set of factors measured grain seed amounts, another field size, and another individual people and animals, for example).

In the last part of the second millennium AD, standardization of factors culminated in the exponential powers of base 10 that we know as the metric system. This standardization of mathematical constructs underlying linguistic forms is not to be confused with standardization of complex numeral (linguistic) expressions discussed by Hurford (1987: 242–284). A change in underlying sequencing strategy would precede standardizations in linguistic form. Linguistic dialect borrowing in numeral systems is thus intertwined with differences in numeral systems as well as linguistic forms (cf. Luján 1999). Similarly, etymologies of the IE digits now need to be reviewed in terms of stages in the evolution of strategies underlying counting (Carruba 1999).

Roman numeral notation suggests a system of repeated factor intervals 5 and 2:

Factor Intervals 5 and 2 Sequencing Roman Numerals

5	2	5	2	5	2
— V	— X	— L	— C	— D	— M
<i>quinque</i>	<i>decem</i>	<i>quinqā-</i> <i>gintā</i>	<i>centum</i>	<i>quin-</i> <i>genti</i>	<i>mille</i>

Roman numeral V is defined by the factor interval 5, X by the factor interval 2, L by another 5 interval, C by another 2 interval and so on: repeated factor intervals 5 and 2. This system is comparable in type to the sexagesimal system in that both consist of collection units sequenced by irregular but repeated factor intervals. In the sexagesimal system they were intervals 6 and 10 as opposed to Roman numeral intervals 5 and 2. Roman numeral factor intervals, like the symbols themselves, may have been borrowed from Etruscan. If so, they undoubtedly originated in an economic function for which factors 5 and 2 were particularly handy.

3.4 Subdividing the factor intervals

The value of the Roman numeral evidence lies less in the symbolic forms of the numerals or their particular set of factor intervals than in what they demonstrate about the way of thinking about the nature of the sequenced units. Roman numeral units V, X, L, C, D, and M have discrete forms that have meanings

derived from positions in the sequence. Intervals between them are subdivided by digit-like strokes. These strokes, units counterpart to digits, are, however, not sequenced. No individual stroke has a meaning any different from another, except as it is positioned beside a member of the sequenced set. Recall first the notation strategy:

Sequenced and Unsequenced Units

I II III IV V VI VII VIII IX X
 X XI XII XIII XIV XV XVI XVII XVIII XIX XX
 ... xL L Lx Lxx ... xC C CC CCC ... cM M CM ...

Given that symbols V, X, L, C, D, and M are the sequenced units, stroke symbols between them are defined in relation to upper and lower members of the sequence. The formal identity among strokes separates them from sequenced symbols such as V and X and suggests a different mathematical role. Up to three strokes are added to a lower member of the sequenced set, while a subtractive strategy applies just before the high end symbol of each interval. So three strokes are added to V or X, but IV is subtractive just before V, and IX is subtractive before X. In Roman numeral thinking the collection units V, X, L, C, D, and M were the sequenced units. Subdivisions corresponding to decimal digits, teens, decades, or hundreds were not sequenced but formed iconically in relation to sequenced collection units. Digits, teens, and decades were thought of in terms of how many iconically matched units they were beyond the lower unit or how many before the upper.

Calculation of intervals between what Hurford (1987: 240–241) called ‘higher numbers’ are usually additive (e.g., teens) or multiplicative (e.g., decades), but may also be subtractive or based on division. Ifrah (1981: 145–151; 337–347) pointed to the inconvenience of additive notation strategies, while Menninger (1977 [1958]: 74–76) noted that strategies based on subtraction and division, ‘back counting’ and ‘over counting’, are rarer than additive or multiplicative strategies. But no scholar has explicitly distinguished the sequencing of digits as opposed to collection units as a typological way of separating stages of numeral evolution. Yet this is the difference between Roman numeral notation and Latin number words. Repeated strokes of the same shape express the counterpart of digits, while symbols referring to collection units are discrete (V, X, D, C, L, and M) as they are the carriers of sequential meaning.

Repetition of units is, of course, not unique to Roman numeral notation nor to Indo-European notations. As an iconic strategy it is unlikely to be peculiar to one system or group of systems, but when it occurs, it undoubtedly reflects an evolutionary stage of number systems in which a class of units is unsequenced.

Early notations often plotted the distance between one unit and the next by additive repetitions of the previous sign (Nissen et al. 1990: 175–180; Ifrah 1981: 145–156), and cuneiform numeral notations of literate periods used both additive repetitions of a lower sign and the sign ‘lal’ for subtractive strategies (Menninger 1977 [1958]: 76). Such systems of notation suggest that units were not sequenced from one to infinity but that units were sequenced from interval to interval as time is from hour to hour. A digital clock leads us to think of minutes sequentially. Analog clocks, however, still reckon the minutes of an hour as subdivisions of the larger unit, as in expressions such as ‘quarter past 2’ and ‘half past 2’ that reckon relations after the last interval or ‘quarter till 3’ that reckons in relation to the upcoming interval unit. Hours are the sequenced units, similar to sexagesimal and Roman numeral collection units, with subdivided intervals calculated both additively and subtractively. There are 24 hours in a day, not 1440 minutes, as hour intervals are the sequenced units.

3.5 *The evolution of numeracy*

The Ancient Near East provides data for stages of numeral evolution over ten millennia of continuous development.³ They begin in the eighth millennium with the pre-mathematical token system, continue into the late fourth and early third millennium BC with protoliterate tablets, and culminate in the third millennium BC sexagesimal system of abstract numerals. The protoliterate tablets show that arbitrary sets of factors had begun to sequence impressed token symbols as pre-numerical quantity units by the third millennium. The Sumerian system of abstract sexagesimal numerals represents a standardization of less regular protoliterate measure-like systems of counting:

Chronology of the Artifacts

8th — Late 4th m. BC	Token Systems	concrete counting
Late 4th — Early 3rd m. BC	Systems of Numeration	pre-numerical counting
Mid-3rd m. BC:	Sumerian Sexagesimals	abstract counting

The data of these successive systems attest the evolution of writing and counting as an interactive process. Symbolic shapes for commodities have counterparts in the symbols of a script that was used to record the economic details of Mesopotamian life as cities arose and trade developed.

3.5.1 *Tokens*

The tokens of the Neolithic were used for accounting purposes from the Persian Gulf to the Mediterranean and into the Anatolian headwaters of the Tigris and

Euphrates Rivers (Schmandt-Besserat 1992: 33; 35–92; 1996: 15–27). Each different shape was used in one-to-one correspondence to count, as mathematicians say, concretely: one token was matched to one unit of a particular commodity, thus fusing the notions of quantity and commodity in a single symbol. One shape might refer to a bowl of grain, another to a jug of oil, and others to containers of beer or units of sheep (Schmandt-Besserat 1992: 189–190; 1996: 115–117).

By the late fourth millennium BC these three-dimensional artifacts were impressed onto clay tablets beside incised drawings or pictographs (Schmandt-Besserat 1992: 143–149; 192; 1996: 72; 74; 118). The transfer of three-dimensional shapes to two-dimensional surfaces resulted in the invention of writing and a major rethinking of the mechanisms of accounting. Impressed onto a two-dimensional clay surface, the quantity and commodity reference of the token was split between the token impression and an incised drawing. The impression of a token and its drawing were now free to evolve in separate directions and impressions evolved as signs for quantity units (Schmandt-Besserat 1992: 140–141; 1996: 69–70; Nissen et al. 1990: 55–60; 76–89).

3.5.2 *Protoliterate factors*

The Swedish mathematician, Jøran Friberg (1979), noticed that totals on the backs of impressed tablets often valued a sphere as different numbers of cones depending on the pictographic commodity sign associated with cone and sphere impressions. From studying the patterns and totals, he hypothesized that different sets of factor intervals defined the meanings of shapes that were used to quantify different kinds of commodities. In fact values defined by different factor intervals also turned out to be specialized for different commodity types much as the value of the foot is for length but the quart for capacity (cf. Nissen et al. 1990: 175–176).

Whereas tokens came in hundreds of shapes (Schmandt-Besserat 1996: 127–157), their impressions used in systems of numeration are limited to 60 shapes (Nissen et al. 1990: 62). Subsets of these 60 shapes then came to function as sequenced units for quantifying different commodities. Damerow and Englund called the sequenced impressions catalogued as N-1, ..., N-60 ‘number signs’, not numerals, because, although they are sequenced by numerical factor intervals, they do not yet have abstract numeral values (Damerow & Englund 1987, 1989; Nissen et al. 1990: 64–65).

3.5.3 *Medieval Monetary Factors*

The most immediate modern analogy to a protoliterate system of numeration is the traditional British pound system with factors 4, 12, and 20 sequencing non-

numerical units, farthing, pence, shilling, and pound. Like e.g., N-14 of the protoliterate systems, the shilling does not refer to an abstract 12 or 48 but depends for its value on its rank order in a monetary sequence, e.g., 12 pence, but 48 farthings.

Pence, Shillings, and Pounds

Factors:	4	12	20	
	farthing	pence	shilling	pound
	"1"	"4"	"48" / "12"	"480" / "120" / "20"

Unlike Sumerian sexagesimal units *u*, *ñesh*, and *shár* of the third millennium BC, units pence, shilling, and pound quantify one particular kind of commodity, money, and are not abstract numerals.

3.5.4 Interpretation

The value of the Ancient Near Eastern data lies in the fact that they attest continuous stages in the evolution of counting from pre-mathematical token counters and the pre-numerical protoliterate systems of numeration to third millennium sexagesimal collection units.

(17) Stages of Numeral Development

SYSTEM	NUMEROSITY	SEQUENCING	ARTIFACT	FORM
I. Preliterate	pre-mathematical	(one-to-one)	token counter	(cone, sphere)
II. Protoliterate	pre-numerical	collection units	impressed sign	(number signs)
III. Sexagesimal	pre-exponential	collection units	cuneiform sign	<i>u</i> , <i>ñesh</i>
IV. Decimal	exponential	powers	Arabic numeral	"ten, hundred"
I-IV. IE system	changing	changing	changing	*dék'm, *k'mtóm

In preliterate and protoliterate stages we know the units only by geometric forms and their uses. The earliest tablets are called 'protoliterate' because decipherment depended, not on deciphering a linguistic system, but on deciphering an accounting system. With the onset of literacy, word forms are matched to artifacts referring to numeral functions. For Proto-IE we have word forms; artifacts take different forms at different stages in IE language development. Any link between an attested form and its numerosity has been neglected in the emphasis on the links between earlier and later phonological word forms, e.g., the sound correspondences that link *dék'm to *decem*, *taihun*, and 'ten'. The leap in numerosity from Gothic *taihun* to English 'ten', for example, is a leap from a collection unit to a power. With the continuity of attestation in the Ancient Near East we have a measuring stick of ordered changes in numeral systems. Applying this new knowledge to the IE situation now reveals correspondences in numerosity that the old PIE decimal hypothesis masked.

4. Constraining the IE decimal hypothesis

The invention of writing (and counting) in the third millennium BC at Uruk (Powell 1981, 1995) comes millennia after the common IE homeland of the fifth to fourth millennium BC. When the citizens of Sumerian Uruk were developing mechanisms for recording agricultural surpluses for storage, distribution, and trade, the prehistoric Proto-Indo-Europeans were probably matching body parts in one-to-one correspondence with units of goods, as they do now in New Guinea today (Biersack 1982; Comrie 1999). If the Proto-Indo-Europeans were precocious, they may have had tokens with names like *dék'm and *k'mtóm.⁴ By the second millennium BC, when Hittite enters history with the cuneiform script and Mycenaean Greek with Linear B, each has adopted a decimal system, or at least a system with repeated factors of 10. The Romans of the first millennium BC, however, still used a notational system that sequenced units perhaps by repeated factors 5 and 2 and measured intervals between them in unsequenced, iconic units:

Subtractive 9 and 19

5 2 5
... V VI VII VIII IX X XI XII XIII ... XIX X X ... L

Latin and Greek teen formations in fact linguistically preserve such additive and subtractive strategies relative to the upper unit '20' (Latin *vii-gintii*, Greek εἴκοσι) and the lower unit '10' (Latin *decem*, Greek δέκα). Early teen words add now sequenced digit words to 'ten', while the last two teen words subtract them from 'twenty':

Additive & Subtractive Teens

	11	12	13	14	15
Latin	<i>un-decim</i>	DECIM ET DUO	DECIM TRIS-QUE	<i>quattuor-decim</i>	<i>quin-decim</i>
Greek	έν-δέκα	δῶ-δέκα	τρεῖς καὶ δέκα	τέτταρες καὶ δέκα	πέντεκαὶδέκα
	1 + 10	2 + 10	3 + 10	4 + 10	5 + 10
	16	17	18	19	
Latin	<i>se-decim</i>	<i>septem-decim</i>	DUO DEEVIIGINTII	UUN-DEE-VIIGINTII	
Greek	ἕξ-καί-δεκα	ἑπτα-καί-δεκα	δουὶν δέοντες εἴκοσι	ἕνος δέοντες εἴκοσι	
	6 + 10	7 + 10	2 – 20	1 – 20	

At the lower end Latin 12–13 preserve the overt additive strategy with “and” (*et*, *-que*) in *decim et duo* or *decem tris-que*, and Greek uses καὶ δέκα “and ten” all the way to “17”. At the upper end Latin *duo dee-viigintii* “2 from 20”, *uun-dee-*

viigintii “1 from 20” and Greek *δυοῦν δέοντες εἴκοσι* “2 lacking from 20” and *ἕνος δέοντες εἴκοσι* “1 lacking 20” are subtractive.

Bonfante & Bonfante (1983: 79 with note 69), who note that the Romans borrowed this system from the Etruscans, suggest that subtraction as a mechanism is confined to Etruscan and Latin. Menninger’s (1977 [1958]: 74–79) data on this process, however, include examples from Germanic, Sanskrit, and Babylonian. The related process, which Menninger calls ‘over counting’, occurs in Germanic, Celtic, Slavic, and Mayan. Additive and subtractive strategies, including back counting and over counting, are all well suited to calculation with an abacus and need not be an index of contact rather than calculation type. Yet Menninger (1977 [1958]: 76–80) and others who describe the non-sequential cutting of decade intervals as ‘over counting’ order it with ‘back counting’ as a curiosity of some numeral systems. They do not see it as a stage in the development of numeral systems which this study finds natural.

The Roman numeral system with collection units sequenced by factor intervals used additive and subtractive (back counting) strategies for forming all units between its collection units. The early Germanic long hundred system was also based on sequenced collection units. If back counting reflects thinking in which collection units were the sequenced units and the intervals between them were subdivided, then we might expect the formation of Germanic decades, as units between collection units, to betray back counting strategies. Given that the long hundred sequenced collection units by factor intervals 10 and 12, when we then find the Germanic lower decades formed with reflexes of **tigu*, we may conclude that **tigu* was the lower collection unit, that lower decades formed with **tigu* used an additive strategy.

Lower decades built with a **TIGU* or ‘ten’ form

	20	30	60
Gothic	<i>twaim tīgum</i> (dat.)	<i>thrins tīguns</i> (acc.)	<i>saihs tīgum</i> (dat.)
Old High German	<i>zwein-zug</i>	<i>driz(z)ug</i>	<i>sēhs-zug</i>
Old Saxon	<i>twen-tig</i>	<i>thrittig</i>	<i>sex-tig</i>
Middle English	<i>twen-ti</i>	<i>thritti</i>	<i>sex-ti</i>
Old Icelandic	<i>tot-togo</i>	<i>thrír tígir</i>	<i>sex-tígir</i>
	<i>tví-tíán</i>	<i>thriá-tígi</i>	<i>sex-røthr⁵</i>
Icelandic	<i>tut-tugu</i>	<i>thrír tígir; thrjá-tíu</i>	<i>sex tugir; sex-tíu</i>

This strategy is familiar and needs no explanation except as the pattern varies, which it does in Old Icelandic and (modern) Icelandic. Beside **tigu* reflexes, Old Icelandic had the new *-tíán* ‘ten’ and (modern) Icelandic generalizes its ‘ten’ form *-tíu* in decades 30–100 beside forms reflecting **tigu*. The new ‘10’ form

based on sequencing powers is what Szemerényi's decimal hypothesis expected, while the older form, we now realize, was a collection unit.

A reflex of *hund preserving reference to the old collection unit occurs either instead of *tigu or in addition to it in the upper decades:

Upper decades built with variant strategies

	70	80	90	10-ty / 100	[11-ty]	"hundred"
Gothic	<i>sibun-tehund</i>	<i>ahtau-tehund</i>	<i>niun-tehund</i>	<i>taihun-tehund</i> ⁶		<i>hund</i>
OHG	<i>sibun-zug,</i> <i>sibun-zo</i> ⁷	<i>ahto-zug,</i> <i>ahto-zo</i>	<i>niun-zug</i>	<i>zēhan-zug,</i> <i>zēhan-zo</i>		<i>hunt</i>
OSaxon	<i>a(n)t-sibun-ta</i>	<i>ahto-da,</i> <i>ahto-doch</i>	<i>nigon-da,</i> <i>nichon-te</i>			<i>hund</i>
OEnglish	<i>hund-seofon-tig</i>	<i>hund-eahta-tig</i>	<i>hund-nigon-tig</i>	<i>hund-teon-tig,</i> <i>hund(red)</i>		<i>hund-twelf-tig</i>
Oicel.	<i>siau-tigir</i>	<i>átta tiger</i>	<i>núu tigir</i>	<i>tío tiger</i>	[<i>ellefo tiger</i>]	<i>hundrath</i>
Icelandic	<i>sjö tugir</i>	<i>átta tugir,</i> <i>átta-tíu</i>	<i>núu tugir,</i> <i>núu-tíu</i>	<i>tíu-tíu</i>		

Lühr (1977) in fact suggested that Gothic *taihun-te-hund* was 'ten till hund' analyzing the much-disputed *te* as "from, to, till" representing a subtractive strategy. If Germanic *tigu and *hund were collection units similar to the sequenced Roman numeral units, we would not be surprised to find digit-like units added to *tigu reflexes and subtracted from *hund reflexes. Similar to the Roman numeral notation in general and Latin and Greek teen formations in particular, the Germanic decade forms measure the interval between a lower collection unit *tigu and an upper collection unit *hund. Forms with both *tigu and *hund reflexes show *tigu spreading to the upper decades before the loss of old *hund reflexes.

Additive and Subtractive Decades

1 2 3 4 5 6 7 8 9 ⁸TIG 2 3 4 5 6 / 7 8 9 HUND 11 HUND
split in strategy: **sex-TIG** / **ANT-sibun**^a

Taking Old Saxon *sex-tig* and *ant-sibun* as examples, *sex-tig* would be "six" of the *tig* unit added to *tig* and *ant-sibun* would be "seven" on the way to (i.e., "from, before") the *ant* (*hund) unit.⁹ Similarly Gothic *sibun-te-hund* "70" is the "seven on the way to *hund*", a subunit between *tigu* and *hund* closer to *hund*. Old English referencing both collection units in its upper decades, e.g., *hund-seofon-tig* "(from) *hund* 'seven' (after) *tig*" shows *tig* spreading to the upper decades beside *hund* before replacing it. Again the strategy is comparable to that for telling time. The type preserved in English "thir-ty" is the "15 past" type, while Gothic *sibun-te-hund* corresponds to "15 till" in type of strategy. Digital time has now generalized the "15 past" strategy with both "one-fifteen" AND

“one-forty-five”. Balto-Slavic has similarly generalized an additive pattern throughout its decades, but Old Irish *-che / -cho / -go*, Greek *-κοντ-*, Latin *-gint-*, and Armenian *-san / -sun* generalized the subtractive “15 till” strategy.

Szemerényi supported his assumption that the IE decades were based on decimal logic by plausible sound correspondences. Ignoring the role that otherwise anomalous “back counting” and “over counting” strategies had in early IE numeral formations, his phonological solutions have been allowed to mask important evidence for numeral evolution. Understanding now that early numeral systems first sequenced collection units before sequencing digits and powers of a base explains Roman numeral notation, the anomalous teen formations in Latin and Greek, and problems of the Germanic decades. For Proto-IE this means that old forms, **dék'm* and **k'mtóm*, were separate words, words referring to smaller and larger collection units. Decades arose from initial subdivision of intervals between upper and lower collection units by both additive and subtractive strategies. Digits, as subdivisions of sequenced collection units, might be added to one **dék'm* and subtracted from two **dék'm's*, or they might refer to the number of **dék'm's* after the first or on the way to the next collection unit, **hund*. Germanic decade names preserve this reckoning in terms of the interval between lower collection unit **dék'm* and upper unit **k'mtóm*. As the system evolved, the name of one unit was generalized throughout the decades. In Latin it was the upper unit **k'mtóm*, in Lithuanian the lower **dék'm(t)*. The decades are thus independent developments after the breakup, as Germanic most clearly shows.

Acknowledgments

Comments and discussion from the audiences at Austin's University of Texas Linguistic Circle, organized by Mark Louden and Mark Southern; San Antonio's University of Texas Language Research Colloquium organized by Victor Villa; and at the Ninth UCLA Indo-European Conference (May 1997) greatly clarified many points, but I am, of course, fully responsible for the results. The University of Texas' resources, including those of the Linguistics Research Center, made preparation of this paper possible.

Author's address

Carol F. Justus
Center for Middle Eastern Studies
University of Texas at Austin
Austin, TX 78712-9400, USA
cjustus@mail.utexas.edu

Notes

1. These forms reflect relevant language chapters in Gvozdanović (1992) and Szemerényi's discussion of decade problems (1960: e.g., 40–42), the ordinals (70–84; 92–95; 111), and the system generally (141–142).
2. Schmidt (1890: 45–49) understood the relationship between early counting and measuring, but lacked the most recent data on the evolution of the Sumerian sexagesimal system as a standardization of even earlier, less regular systems of weights and measures (Powell 1989).
3. While counting may have developed differently elsewhere, only the Ancient Near East preserves a continuously attested tradition.
4. Vyacheslav V. Ivanov (p.c.) has called attention to token finds lying unstudied in a Hungarian museum and the studies begun by Gimbutas and Winn of an Old European writing system that remain incomplete.
5. This Old Icelandic variant might be compared to “hund-red” below where both *-røthr* and *-red* come from PGmc *RATHA “number”; cf. Latin *reor* “think, calculate, count” (cf. Lehmann 1986: 282 R10).
6. The variant, *taihun-taihund*, is the source of evidence that Germanic “hundred” was ten ten’s; cf. also Icelandic *tíu-tíu* “ten-ten”.
7. Ross & Berns (Gvozdanović 1992: 609–611) take the Old High German suffix *-zo* as reflex of *TOX (a *TEXAN variant?) which Justus (1996: 72) hypothesizes may preserve a third collection unit name.
8. Justus (1996: 68–71) takes this form as related to Crimean Gothic *stega* “20”, not an alternative “10”.
9. Roman numeral use of unsequenced strokes and Latin teen formations differ slightly in that “2 from 20” would have been more “8-viigintii” versus “7-decim” to parallel the Germanic upper decades. This difference represents another step in numeral development yet to be investigated.

References

- Araujo, Frank P. 1975. “Counting Sheep in Basque”. *Anthropological Linguistics* 17:4.139–145.
- Biersack, Aletta. 1982. “The Logic of Misplaced Concreteness: Paiela body counting and the nature of the primitive mind”. *American Anthropologist* 84.811–829.
- Bonfante, Giuliano & Larissa Bonfante. 1983. *The Etruscan Language: An introduction*. New York: New York University Press.
- Comrie, Bernard. 1999. “Haruai Numerals and their Implications for the History and Typology of Numeral Systems”. *Numeral Types and Changes Worldwide* ed. by Jadranka Gvozdanović, 81–94. Berlin & New York: Mouton de Gruyter.
- Carruba, Onofrio. 1999. “Die indogermanischen Zahlwörter: Neue Ergebnisse und Perspektiven”. *Language Change and Typological Variation: In honor of Winfred P. Lehmann*. Vol. I: *Language Change and Phonology*, 145–171. ed. by Edgar C. Polomé & Carol F. Justus. (Journal of Indo-European Studies Monograph, 30.)

- Damerow, Peter & Robert K. Englund. 1987. "Die Zahlzeichensysteme der Archaischen Texte aus Uruk". *Zeichenliste der archaischen Texte aus Uruk* ed. by Margaret W. Green & Hans J. Nissen, 117–166. (= Ausgrabungen der deutschen Forschungsgemeinschaft in Uruk-Warka, 11. Archaische Texte aus Uruk, 2.) Berlin: Gebrüder Mann.
- Damerow, Peter & Robert K. Englund. 1989 "The Proto-Elamite Texts from Tepe Yahya". *Bulletin of the American School of Prehistoric Research* 39. Peabody Museum, Harvard University.
- Danzig, Tobias. 1959. *Number: The language of science*. 4th ed. New York: Macmillan.
- Friberg, Jøran. 1979. "The Early Roots of Babylonian Mathematics: II. Metrological relations in a group of semi-pictographic tablets of the Jemdet Nasr type, probably from Uruk-Warka". (= *Research Report*, 1979–15.) University of Göteborg. Department of Mathematics, Chalmers University of Technology.
- Gamkrelidze, Thomas V. & Vyacheslav V. Ivanov. 1983. "The Ancient Near East and the Indo-European Question: Temporal and territorial characteristics of Proto-Indo-European based on linguistic and historico-cultural data". *The Indo-European Problem: Recent Soviet contributions* ed. by Marvin A. Powell, Jr., 7–52 (= *Soviet Studies in History* XXII. 1–2).
- Gvozdanović, Jadranka, ed. 1992. *Indo-European Numerals*. New York: Mouton de Gruyter.
- Høyrup, Jens. 1994. *In Measure, Number, and Weight: Studies in mathematics and culture*. Albany: State University of New York Press.
- Hurford, James R. 1987. *Language and Number: The emergence of a cognitive system*. New York: Basil Blackwell.
- Ifrah, Georges. 1981. *Histoire universelle des chiffres*. Paris: Seghers.
- Justus, Carol F. 1996. "Numeracy and the Germanic Upper Decades". *Journal of Indo-European Studies* 24.45–80.
- Lehmann, Winfred P. 1986. *A Gothic Etymological Dictionary*. Based on the 3rd ed. of Sigmund Feist's *Vergleichendes Wörterbuch der Gotischen Sprache*. Leiden: Brill.
- Luján Martínez, Eugenio Ramon. 1999. "Towards a typology of change in numeral systems". *Language Change and Typological Variation: In honor of Winfred P. Lehmann*. Vol. I: *Language Change and Phonology*, 183–200. ed. by Edgar C. Polomé & Carol F. Justus. (Journal of Indo-European Studies Monograph, 30.)
- Lühr, Rosemarie. 1977. "Die Dekaden '70–120' im Germanischen". *Münchener Studien zur Sprachwissenschaft* 36.59–72.
- Mallory, James P. 1989. *In Search of the Indo-Europeans. Language, archaeology, and myth*. New York: Thames & Hudson.
- Menninger, Karl. 1977 [1958]. *Number Words and Number Symbols: A cultural history of numbers*. Transl. by Paul Broneer from the rev. German original. Cambridge, Mass.: MIT Press.

- Nissen, Hans J., Peter Damerow & Robert K. Englund. 1990. *Frühe Schrift und Techniken der Wirtschaftsverwaltung im alten Vorderen Orient: Informationsspeicherung und -verarbeitung vor 5000 Jahren*. Berlin: Franzbecker.
- Powell, Marvin A., Jr. 1981. "Three Problems in the History of Cuneiform Writing: Origins, direction of script, literacy". *Visible Language* XV.419–440.
- . 1989. "Maße und Gewichte". *Reallexikon für Assyriologie und Vorderasiatische Archaeologie* ed. by Dietz Otto Edzard, Band 7, Lieferung 5/6:457–480. Berlin & New York: Walter de Gruyter.
- . 1995. "Metrology and Mathematics in Ancient Mesopotamia". *Civilizations of the Ancient Near East* ed. Jack Sasson (Assoc. eds. John Baines, Gary Beckman & Karen S. Rubinson), 1941–1957. (Macmillan Library Reference USA.) New York: Scribner's Sons.
- . 1997. "Sexagesimal System". *Encyclopedia of the History of Science, Technology, and Medicine in Non-Western Cultures* ed. by Helaine Selin, 892–893. Dordrecht, Boston & London: Kluwer.
- Schmandt-Besserat, Denise. 1992. *Before Writing*. Vol. I: *From Counting to Cuneiform*. Austin: University of Texas Press.
- . 1996. *How Writing Came About*. Austin: University of Texas Press.
- Schmidt, Johannes. 1890. *Die Urheimat der Indogermanen und europäische Zahlssystem*. (= Abhandlungen der Königlichen Preussischen Akademie der Wissenschaften zu Berlin.)
- Sommer, Ferdinand. 1951. *Zum Zahlwort*. (= Sitzungsberichte der Bayerischen Akademie der Wissenschaften. Philosophisch-historische Klasse, Jahrgang 1950, Heft 7.) Munich: Bayerischen Akademie der Wissenschaften.
- Szemerényi, Oswald. 1960. *Studies in the Indo-European System of Numerals*. Heidelberg: Carl Winter.
- Ulf-Møller, Jens. 1991. "The Higher Numerals in Early Nordic Texts and the Duodecimal System of Calculation". *The Audience of the Sagas*, 323–330. (= The Eighth International Saga Conference, 2.). Gothenburg University.

CHAPTER 31

Berthold Delbrück and his Contemporaries on 'Tempora' in Sanskrit

Vit Bubenik

Memorial University of Newfoundland

1. Introduction

The last two decades of the previous century saw a sequence of brilliant studies of Sanskrit morphology and syntax: the first edition of W. D. Whitney's *Sanskrit Grammar* in 1879; J. S. Speijer's *Sanskrit Syntax* in 1886; B. Delbrück's *Altindische Syntax* in 1888; the second edition of Whitney's *Sanskrit Grammar* in 1889; Speijer's *Vedische und Sanskrit-Syntax* in 1896; culminating in Delbrück's monumental *Vergleichende Syntax der indogermanischen Sprachen* (1900). Delbrück's *Altindische Syntax* was preceded by several studies of salient morphological and syntactic problems (1867, 1869, 1871, 1878), among them a highly influential study of Sanskrit tenses in 1876, entitled *Altindische Tempuslehre*.

Whitney (1879: v) stated specifically in the preface to the first edition of his grammar that he intends

to make a presentation of the facts of the language primarily as they show themselves in use in the literature, and only secondarily as they are laid down by the native grammarians.

One of the salient features of his presentation is his positivist penchant for statistics which makes his grammar different from that of his predecessors (Bopp, Benfey, Kielhorn, Monnier-Williams, Müller). Statistics, in Whitney's conviction, "help an appreciation of the character of the language"; and it precisely is here where "the native grammar is especially deficient and misleading" (1879: vii).

Speijer's *Sanskrit Syntax* (1886) held the distinction of being the first complete syntax of Classical Sanskrit. It was introduced to the scholarly world by H. Kern (Leyden, 1886) who pointed out that Pāṇini's famous grammar — in spite of being superior in many respects to anything of this sort produced by other "civilized nations of antiquity" — was deficient in its treatment of syntax. In his writing about the matters of Sanskrit tenses Speijer already had an illustrious precursor in Delbrück (1876).

Two years after the appearance of Speijer's monograph the scholarly world could avail itself of two more 'definite' versions of Sanskrit *tempuslehre* in Delbrück's *Altindische Syntax* (1888) and the second edition of Whitney's *Sanskrit Grammar* (1889).

2. Whitney's morphological approach

Whitney's Sanskrit *tempuslehre* (1889/1969: 201) consists of an enumeration of the categories and comments on their morphology and usage. All of them are classified as 'tenses':

1. *present*
2. *imperfect*, "closely related with it in form, having a prefixed augment"
3. *perfect*
4. "a so-called *pluperfect*, made from it with prefixed augment"
5. *aorist* (simple, reduplicated, sigmatic)
6. *conditional*, "an augment tense standing to it in the relation of an imperfect to a present"
7. *future*
8. and *second (periphrastic) future* (not found in the Veda).

The categories in 2, 4, 5, and 7 display so-called 'secondary' suffixes (-*m*, -*s*, -*t*, etc.) while their augment is optional (as in Homeric Greek). In spite of what Whitney said about the formation of the pluperfect, "an augment-preterit from the perfect-stem, to which the name of pluperfect is given on the ground of its formation (though not of its meaning)" (p. 295), he left its slot empty in his table of tense systems (in the Appendix on p. 520). Also no past counterpart was provided to the second future. Omitting the modal and participial forms, Whitney's table is reproduced here as Table 1.

	Present	Perfect	Aorist	Future
Indicative	<i>karómi</i> (1)	<i>cakāra</i> (3)		<i>kariṣyāmi</i> (6) <i>kartāsmi</i> (8)
Augmented preterit	<i>ákaravam</i> (2)		<i>ákārṣam</i> (5)	<i>ákariṣyam</i> (7)

Table 1. Whitney's system of Sanskrit tenses (root *kr* "do, make")

3. Delbrück's semantic approach

Delbrück's *tempuslehre* is much more extensive (1888: 273–301). He opens his treatment by expressing doubts that one can talk about 'basic' meaning (Grundbegriff) of individual subsystems (present, aorist, perfect, and future); this task, however, he will take on later in his comparative syntax of IE languages (1900). A lot of his discussion revolves around the issues of morphology (suppletion), the use of particles (e.g., present + (*purá:*) *ha sma* expressing the preterite), and contextual meanings (presens historicum, future time reference, and modal use of the present).

According to Delbrück (1888: 279), the imperfect is a narrative tense "durch welches der Hörer aufgefordert wird, sich mit seiner Phantasie in die Vergangenheit zu versetzen". Its salient feature is the fact that unlike the aorist and perfect it never possesses any relationship to the present.

To use his fitting example from Śatapatha-Brāhmaṇa (11, 5, 1, 7), when Urvaśī said to Purūravas

- (1) *ná vái tvam̐ tād akaror yád ahám abravam*
 not PRT you it do+IMPF+2SG what I say+IMPF+1SG
 she meant "You didn't (then) do that, what I (had) said"
 i.e., **not** "You haven't done that, what I have said".

Delbrück's ideas à propos the perfect (1876: 101ff) crystallized during his work on the comparative syntax of IE languages. Twenty years later (1897: 269) he thought of the perfect according to the currently accepted definition as the 'resultative' (Comrie 1976) category: "the result reached in the present and accordingly the completed action in the present" (er drückt ... den in der Gegenwart erreichten Zustand und danach die in der Gegenwart vollendete Handlung aus). In his *Altindische Syntax* (pp. 299–301) he discussed the perfect under three headings: the perfect (used) as the present, the perfect as a past tense in non-narrative sense, and the perfect as a narrative tense. The sphere of usage

of the narrative perfect intersects with that of the aorist, and that of the narrative perfect with that of the imperfect. Interesting statistical surveys from the Vedic *saṃhitās* and Brāhmaṇas regarding two types of narration introduced by either *abravīt* (imperfect) "he said" or *uvāca* (perfect) "he said" are presented on p. 300. Alas, they are inconclusive and Delbrück surmised that one ends up with an 'impression' that the imperfect and the narrative (historical) perfect are coextensive (sich völlig gleich stehen). Whitney (1889: 296) was in agreement with Delbrück; in addition, however, he reminded his readers of the Hindu grammatical lore, according to which "the perfect is used in the narration of facts not witnessed by the narrator", but he was quick to qualify his statement by emphasizing that there was no evidence of its being used so ("either exclusively or distinctively").

4. The first monograph on Sanskrit syntax

Three years earlier in his *Sanskrit Syntax* (1886: 247), Whitney's Dutch contemporary J.S. Speijer was much more positive about the Hindu grammatical tradition in pointing out that "the practice of good authors is generally in accordance with this statement". As an example of 'good' literary works Speijer mentioned the famous historical romance *Daśakumāracarita* "The Experiences of Ten Princes" composed by the rhetorician Daṇḍin in the 8th C., and another work entitled *Kathāsaritsāgara* "The Ocean of Stories" written in poetic form by Somadeva in the 11th C. Daṇḍin uses imperfects and aorists 'promiscuously' in passages narrated by the very persons who have experienced them. When, however, the author himself is speaking or when one of his characters relates 'a fable of olden times', the perfects make their appearance side by side with imperfects and aorists. The last point is important because it does not follow that the imperfect is restricted to relating the past events witnessed by the speaker. The imperfect is used for events which are and which are not "within the compass of the speaker's experience", concluded Speijer (p. 297).

It should be mentioned that the Pāṇinian term *parokṣe* (literally beyond the eye) "being out of sight, invisible, imperceptible", delimiting the sphere of usage of the perfect, proved to be in need of an amendment already in antiquity. This was done by an illustrious commentator on Pāṇini, Kātyāyana (3rd C. B.C.), along the following lines (preserved in Patañjali's *Mahābhāṣya* II.119):

well known facts falling within the speaker's sphere of observation are to be put in the imperfect — not in the perfect — even if they have in fact not been witnessed by him.

This observation squares remarkably well with those made à propos the situation in modern languages with the category of 'inferential' (such as Turkish and Bulgarian). These languages use the inferential when expressing the fact that the speakers base their statements on reported information; when, however, narrating historical events not witnessed by them they may revert to their non-inferential categories (imperfect, aorist). Here Kātyāyana heralds modern studies of inferentiality (called various terms such as 'reported information', 'non-confirmation', 'evidentiality', 'interference', etc.).

In Classical Sanskrit, as described by Speijer (1886: 246), the main distinction is that "between such past facts, as have not lost their actuality, and such as have, and therefore belong to history". The latter historical events may be expressed by any aspectual form (aorist, imperfect, perfect, or participle) while the 'actual' events (past events with present relevance) only by the aorist or the participle (in later Sanskrit); or, to define the imperfect and perfect negatively (with Speijer 1886: 247), "they cannot be used except of such facts as have lost their actuality for the speaker".

5. Delbrück's 'tempuslehre'

In his *tempuslehre* (1876: 5ff) Delbrück characterized the Vedic aorist as the category which expresses what from the point of view of the speaker has recently happened; whereas the aorist as used in the Vedic prose expresses what the speaker himself has experienced. Twelve years later (1888: 280) Delbrück admitted that both characterizations are falsifiable and despaired of being able to come up with a completely satisfactory concise term which would cover both usages. Contrasted with the imperfect, the aorist never describes (schildert) but only announces (mittheilt) that a certain event took place (ein Vorgang in die Erscheinung getreten ist). Relying on the state of affairs in Ancient Greek where the aorist is characterized in terms of Aktionsart categories ('momentary' or 'punctual'), Delbrück suggested that in the Sanskrit aorist (aoristische Äusserung) the view of duration is irrelevant (der Gesichtspunkt der Zeitdauer gar nicht in Betracht kommt). What matters is simply that an event took place. Thus, as is well known, the aorist may be used to express events which lasted very different time spans. In his presentation of data (1888: 281–289) Delbrück preferred to subdivide his examples according to the 'situation' in which the aorist occurs: (i) direct speech, i.e., the aorist forms "put into the mouth of speakers" by the writer, and (ii) those where the writer himself is the speaker. This dichotomy, however, is not categorical. Outside direct speech, Delbrück (pp. 285–286)

provided examples of the aorists expressing an event which took place immediately before the moment of utterance (but also in a more distant past) and paid special attention (pp. 287–289) to those situations when the reference point is the ritual activity (to be described or to be explicated). In the latter case there are two options: the aorist expressing what immediately follows the ritual (is contemporary with it), or what has to be thought of as finished before the ritual.

6. Back to the ancient Hindu grammarians

While Whitney failed — and Speijer and Delbrück were apparently not interested — in presenting the Sanskrit *tempuslehre* as a morphological system, the native Hindu grammarians were more successful in this respect. Pāṇini's grammar introduces so-called L-suffixes after the verbal root on two conditions: that an 'agent' is to be denoted (in common with all L-suffixes) and that the action referred to is specified in its 'time'. The present was called *laṭ*, the perfect *liṭ*, the aorist *luṭ*, etc. There were eight such meta-linguistic terms which may be organized as shown in Table 2.

laṭ	liṭ	luṭ	lṛt
laṇ	liṇ	luṇ	lṛṇ
Present	Perfect	Periph. Future	Future
Imperfect	Optative	Aorist	Conditional

Table 2. Sanskrit tense system according to Pāṇini (with modern equivalents)

The above metalinguistic terms may be analyzed as indicating the awareness of:

- (i) the aspectual triad present-aorist-perfect (indicated by the vowel *a* vs. *u* vs. *i*);
- (ii) the distinction of the aspectual forms marked by the primary suffixes (indicated by C2 = *ṭ*) vs. those marked by the secondary suffixes and the augment C2 = *ṇ*.
In contemporary terminology: *ṭ* = non-past, *ṇ* = past;
- (iii) the remaining puzzle is the pairing of the perfect (*liṭ*) with the optative (*liṇ*), where we would like to see the pluperfect instead of the optative;
- (iv) the pair future (*lṛt*): conditional (*lṛṇ*) exactly parallels the pair present (*laṭ*): imperfect (*laṇ*);
- (v) and, finally, it should be observed that Pāṇini does not treat the periphrastic future forms (*luṭ*, i.e., the non-past counterparts to the aorist *luṇ*) as analytic (cf. Rocher 1965).

7. Analytic aspectual forms in late Classical Sanskrit

Point (v) above brings us to the existence of analytic aspectual forms in late Classical Sanskrit which were handled only marginally by Whitney and Speijer (Delbrück's studies were devoted to the Vedic and Brāhmaṇa periods when the analytic forms only made their first appearance). A propos (v) Speijer (1886: 259), in accordance with Pāṇini, characterized the periphrastic future as a 'remote' future, versus the 'general' sigmatic future which may express any future action "whether actual or remote". A fitting example from *Maitrāyaṇī Saṃhitā* was provided by Whitney (1889: 338):

- (2) *adyā varṣiṣyati śvó vraṣṭā*
 today rain+FUT+3SG tomorrow rain+PERIPH FUT
 "It is going to rain today; it will rain tomorrow".

Whitney (1889: 394–395) also provided a brief survey of participial periphrastic formations in the Brāhmaṇas (with occasional instances from Vedic Sanskrit). They consist of the main verb in its participial form accompanied by one of the 'proto-auxiliaries' (my term, VB): *i-* "go", *car-* "go", *ās-* "sit", *sthā-* "stand", *as-* "be" and *bhū-* "become". The former four express the habitual (continuous) aspect; the latter two may form the expressions of the progressive and retrospective aspect (i.e., the perfect, pluperfect, and future perfect "in no regular and continued fashion", p.395). Here are some of his examples: *vibhājann éti* (*go*+PRES) [Rig-Veda] "he ever gives away"; *dahann ait* (*go*+IMPF) [Tāṇḍya-Brāhmaṇa] "he kept burning"; and *prativādato 'tiṣṭhan* (*stand*+IMPF) [Aitareya-Brāhmaṇa] "they kept refusing".

Another late periphrastic formation — virtually absent from Vedic and Brāhmaṇa Sanskrit — is the combination of the activated past passive participle (*kṛta* "done" + *-vant* = *kṛtavant* "having done") with the copula: *kṛtavān asmi* "I have done". Its spread may partly be understood as a 'reaction' against the ergative construction based on the passive construction of the type *mayā kṛtam* literally by-me done > "I have done" (cf. Bubenik 1993 for details).

In late Classical Sanskrit the 'proto-auxiliary' *sthā-* "stand" became the most widely used one for the expression of tense. For instance, in Jambhaladatta's *Vetālapañcaviṃśatikā* "The Twenty-Five Tales of the Demon" (of the 14th C.) the whole paradigmatic set of analytic expressions for the progressive aspect is available. Under periphrastic 'past' one finds not only the analytic auxiliary *tiṣṭhati sma* (stands = PAST) but also the formations such as *kurvans atiṣṭhat* (with the auxiliary in the imperfect) and *kurvans asthāt* (with the auxiliary in the aorist). All these formations are surveyed in Table 3.

Basic forms	Periphrastic counterparts
(Whitney's) Future (6)	<i>kurvans sthāsyati</i> "he will be doing"
Present (1)	<i>kurvans tiṣṭhati</i> "he is doing"
'Past': Imperfect (2), Aorist (5)	<i>kurvans tiṣṭhati sma</i> "he was doing"
Analytic Perfect: <i>tena kṛtam</i> (ergative) ~ <i>kṛtavān</i> (non-ergative)	<i>kurvans tasthau</i> "he has been doing"

Table 3. Analytic aspectual forms in late Classical Sanskrit

8. Towards recognition of grammatical aspect by structuralists

The overall picture of Sanskrit *tempuslehre* emerging from these three 19th C. 'classics' with all their merits discussed above is disappointing in one important point: there are no explicit observations on the expression of aspect and Aktionsart. Whitney's grammar (1879, 1889) does not include any chapter on syntax where these matters should be dealt with. In his monographs on syntax (1886, 1896), Speijer conceived of the Sanskrit 'tempora' more in terms of their actuality which translates into their relative position in the past and the future time zones: Imperfect — **Historical (Far) Past**, Aorist — **Actual (Near) Past**, mirrored by Future — **Near Future**, Periphrastic Future — **Far Future** (cf. 3 and 6). Following the native tradition, the perfect was conceived of as the chief exponent of the 'inferentiality' (with a proviso regarding non-witnessed but well known historical events). Delbrück (1888: 280) came closest to the realization that verb forms possess aspectual functions in his discussion of the aorist (see 4). Further moderate inroads in this direction were made by his followers during the early 20th C. In his *Handbuch des Sanskrit* (1905: 282ff), A. Thumb used the term Aktionsart (in the sense of grammatical aspect) in his characterization of 'temporal stems' (die Tempusstämme nicht die Zeitstufe, sondern die sog. Aktionsart bezeichneten ... wie sich noch deutlich im Griechischen erkennen läßt). In his influential Vedic grammar (1916/1953: 345), A. A. Macdonell echoes Speijer and Delbrück in saying that

the imperfect is the past tense of narration, never having any relation to the present ... the aorist ind. ... neither describes nor indicates duration, but simply states a fact.

Further progress in our understanding of aspectual matters of Sanskrit came after two generations during the era of structuralism. In aspectology an epoch-making event was the study of Greek tenses and aspects by the Spanish scholar M. S.

Ruipérez in 1954. Inspired by this work the Dutch Sanskrit scholar J. Gonda ushered in a new age of Sanskrit morphosyntactic studies with his monograph *The Aspectual Function of the Rgvedic Present and Aorist* (1962).

Author's address

Vit Bubenik
Department of Linguistics
Memorial University of Newfoundland
St. John's, Newfoundland, CANADA A1B 3X9
vbubenik@morgan.ucs.mun.ca

References

- Bubenik, Vit. 1993. "Passives and ergatives in Middle Indo-Aryan". *Historical Linguistics 1993* ed. by H. Andersen, 49–57. Amsterdam: Benjamins.
- Comrie, Bernard. 1976. *Aspect: An introduction to the study of verbal aspect*. Cambridge: Cambridge University Press.
- Delbrück, Berthold. 1867. *Ablativ, localis, instrumentalis im altindischen, lateinischen, griechischen und deutschen*. Berlin: Dümmler.
- . 1869. "Ueber den indogermanischen, speciell den vedischen Dativ". *Zeitschrift für vergleichende Sprachforschung* 18.81–106.
- . 1871. "Der Gebrauch des Conjunctivs und Optativs im Sanskrit und Griechischen". *Syntaktische Forschungen*, 1. Halle: Waisenhau.
- . 1876. "Altindische Tempuslehre". *Syntaktische Forschungen*, 2. Halle: Waisenhau.
- . 1878. "Die altindische Wortfolge aus dem Śatapathabrāhmaṇa dargestellt". *Syntaktische Forschungen*, 3. Halle: Waisenhau.
- . 1888. *Altindische Syntax*. *Syntaktische Forschungen*, 5. Halle: Waisenhau. Repr. 1968. Darmstadt: Wissenschaftliche Buchgesellschaft.
- . 1900. *Vergleichende Syntax der indogermanischen Sprachen*. (= Vol. 5 of K. Brugmann & B. Delbrück's *Grundriß der vergleichenden Grammatik der indogermanischen Sprachen*). Straßburg: Trübner.
- Gonda, Jan. 1962. *The Aspectual Function of the Rgvedic Present and Aorist*. The Hague: Mouton.
- Hewson, John & Vit Bubenik. 1997. *Tense and Aspect in Indo-European Languages. Theory, Typology, Diachrony*. Amsterdam & Philadelphia: Benjamins.
- Macdonell, Arthur A. 1916/1953. *A Vedic Grammar for Students*. Oxford: Oxford University Press.

- Ruipérez, M. S. 1954. *Estructura del sistema de aspectos y tiempos del verbo griego antiguo. Análisis funcional sincrónico*. Salamanca. [*Structure du système des aspects et des temps du verbe en grec ancien. Analyse fonctionnelle synchronique*. Annales littéraires de l'Université de Besançon, Les Belles Lettres, 1979.]
- Rocher, Rosane. 1965. "La formation du futur périphrastique sanskrit selon Pāṇini: un exemple de description linguistique". *Annali, Istituto Orientale di Napoli* [AIÖN], *sezione linguistica* 6.15–22.
- Speijer [= Speyer], J. S. 1886. *Sanskrit Syntax*. Leiden: Brill. (Repr. 1973, Delhi: Motilal Banarsidass).
- . 1896. *Vedische und Sanskrit-Syntax*. (Grundriss der Indo-Arischen Philologie und Altertumskunde, 1:6.) Strassburg: Trübner.
- Thumb, Albert. 1905. *Handbuch des Sanskrit*, I. Heidelberg: Winter.
- Whitney, William Dwight. 1879. *Sanskrit Grammar*. 1st ed. [*Indische Grammatik* aus dem Engl. übersetzt von Heinrich Zimmer.] Leipzig.
- . 1889. *Sanskrit Grammar*. 2nd ed. Cambridge, Mass.: Harvard University Press. [2nd Indian Edition 1969, Delhi: Motilal Banarsidass].

PART VI

Latin and Comparative Romance Linguistics

CHAPTER 32

Observations on Two Recently Discovered Latin Inscriptions

Philip Baldi

Pennsylvania State University

The modern history of Indo-European scholarship is well-endowed with examples of discoveries which have reshaped conceptions of linguistic history. Some of these events, such as the discovery of the trove of Hittite materials uncovered in the royal archives of Hattušaš at Boğazköy, or the discovery of the Tocharian manuscripts in Chinese Turkestan, or the decipherment of the Linear B script and the establishment of the language it records as Mycenaean Greek, have truly redirected the course of Indo-European linguistics in this century. The discovery and classification of Hittite as an IE language, for example, led to the eventual confirmation of Saussure's PIE "coefficients sonantiques", securing, at least to many people's satisfaction, their existence in the protolanguage. The discovery of the centum Tocharian so far to the east of the general IE-speaking area brought about a redesign of the IE dialect map. And many of the details of the deciphered Linear B script led to the uncovering of features in Mycenaean which had been conjectured on the basis of comparative and internal reconstruction. For example, Mycenaean more faithfully reflects inherited labiovelars in environments where other dialects of Greek merged them with labial and dental stops: cf. Gk. τε "and", Myc. *-qe* < **-k₂e*; Gk. βοῦς "cow", Myc. *qo-u-ko-ro* "cow-herd" < **g₂ou-*. For the most part these discoveries have had a comforting confirmatory effect on IE scholarship, verifying hypotheses, stabilizing reconstructions, and validating previous conjectures.

Of course, events like the discovery of a new language from the IE family are relatively rare. Far more common are the events which pass by unnoticed by the general community of linguists, but which are of enormous significance to those who work on the minute details of individual languages and try to fit them into the overall IE picture. I am thinking here of things like the recent finds of

Continental Celtic such as the Botorrita text; the new leaf of the Gothic Bible; the dating of the Old Prussian Basel Epigram as the oldest document attesting the Baltic languages; and numerous other items of detail which keep the discipline of Indo-European linguistics from empirical stagnation.

In the past two decades there have been two important finds which have contributed significantly to the understanding of the history of Latin and the Italic languages, both of which have found, or soon will find, their way into the general Indo-Europeanist canon. The first of these was the discovery of the Lapis Satricanus, which furnished rich material for Latin epigraphy and the database of the preliterary language. The Lapis Satricanus is a two line example of archaic Latin which dates to the period just before 500 BCE. It was discovered in the fall of 1977 during an archaeological investigation at the ancient city of Satricum, about 60 kilometers to the south and slightly to the east of Rome. The inscription is incised on stones that formed part of the substructure of the temple of Mater Matuta which had been covered over by growth and debris. Though there is debate about the essential Latin character of the inscription, with some scholars seeing Oscan, Faliscan, or Volscian features, its essentially Latin character seems as secure as the alternatives. Translations of the inscription generally agree on the division and meaning of all the words beginning with *POPLIOSIO*.

CIL I².2832a; Stibbe et al. (1980), De Simone (1980), Campanile (1985), Prosdocimi (1985), Wachter (1987: 75–80), Lejeune (1989), Silvestri (1993).

IEISTETERAI POPLIOSIO VALESIOSIO SVODALES MAMARTEI

The form *IEI* before *STETERAI* is probably the last part of a word which has been partially destroyed on the stone, perhaps [*SAL*]*IEI* = [*Sal*]*i* “The Salic (priesthood or fraternity)”, suggesting the following epigraphical representation and translation:

SAL]*IEI* STETERAI POPLIOSIO VALESIOSIO SVODALES MAMARTEI

“Members of the Salic fraternity of Publius Valerius have erected [this] to Mars”.

STETERAI is probably a 3rd perf. pl., corresponding roughly to the Latin perfect ending *-ēre*, though this analysis is not without its problems, since the final *-erai* of *STETERAI* would correspond to a later Latin **-ērī*. Other matters of more general linguistic interest are the old genitives in *-osio* in *POPLIOSIO* and *VALESIOSIO*, a case ending previously unattested in Latin, at least in this form (it may be concealed in genitives like *eius* “his, hers” and *cuius* “whose”; see Watkins 1966: 38). Both Italic and Celtic typically mark the genitive singular of *o*-stem

nouns with *-ī*. Representative forms are Ogham Ir. *maqī* "of the son", Lat. *lupī* "of the wolf". However, Faliscan preserves the form **-osjō*, e.g., **eco quon Evotenosio** "I am the vase of Evotenos" (Vetter 1953: 285, no. 242 B); cf. Skt. gen. sg. *tásya* "his", *devásya* "of the god", etc., so the presence of this suffix in the Lapis Satricanus inscription strengthens its role in reconstructed PIE generally. The unrhottacized *s*'s in the **POPLIOSIO** and **VALESIOSIO** show that the rule by which Latin intervocalic **s* was transformed into *-r-* had not yet applied, preserving the older forms of *Publiū* and *Valeriū*. The form **SVODALES** "members of a fraternity" is particularly interesting. The traditional etymology of this word (CLat. *sodālis*) has it derived from a putative **suodālis*, which according to established sound laws would yield *sodālis*. The PIE background of the word is **sue-d^h(e)h₁-* "own established usage, custom". The SV- of the Lapis Satricanus thus confirms the traditional etymology, achieved by comparative and internal analysis (Watkins 1989: 786–789). And finally this inscription contains the dative ending *-ei* (later *-ī*) in the unusual reduplicated theonym **MAMARTEI**. This ending, reflecting PIE **-ei*, is found in other inscriptional examples such as **VIRTUTEI** = *virtūtī* "to valor" and **APOLONEI** = *Apollōnī* "to Apollo", but has been replaced by the time of the later literary documents on which we largely rely.

So it can be seen that the discovery and interpretation of this brief, five or six word inscription of archaic Latin has not only had an effect on our understanding of some details of the history of Latin, but has also contributed to the picture of the relationship of Latin to other Italic languages, and other members of the IE family.

A second, very recent addition to the empirical base of the history of Latin has been announced and published by Cristofani (1996). This is the inscription on the "Garigliano Bowl", a small "scodella d'impasto" which apparently comes from the area of a sanctuary of the goddess Marica near the mouth of the Garigliano river, between Lazio and Campania, where excavations were first undertaken in the 1920s. The sanctuary lies in the ancient territory occupied by the Aurunci, between the Latini and the Volsci. Though the details of the discovery of this bowl are unclear (it seems to have been in private hands for some time), the archeological, paleographical, and linguistic issues are now openly accessible following the publication of Cristofani (1996), which will serve as the definitive departure point for any future discussion involving this find.

The inscription, which has been dated by Cristofani to the first half of the 5th C. BCE, is written in a variant of the Western Greek/Etruscan alphabet. There are several unclear letters on which the divergent readings discussed below are centered. The inscription is written on two parts of the bowl. The first (A. below) contains a single name, apparently in a non-Latin language. The

second part (B. below) is much longer, and contains several interesting archaic forms. B. is written without word-breaks in a continuous circle. Both parts are read left-to-right. As interpreted by Cristofani, the inscription runs as follows:

A. αηυιδιες

B. παριμεδεσομκομμεουσσοκιοιστριφοιαδδεομδυο[—]νει

With word breaks:

παρι μεδ εσομ κομ μεουσ σοκιοισ τριφοιαδ δεομ δυο[—]νει

Mancini (1997) reorders the first two words *παρι μεδ* to form a syntactic unit with *νει*, reads the final *δ* of *τριφοιαδ* as an *ι* (admitted as a possibility by Cristofani), and conjectures [*ναι*] after *δυο*, with the following (transliterated) interpretation:

A. **AHUIDIES**

“Avidius”

B. **ESOM KOM MEOIS SOKIOIS TRIVOIAT DEOM DUONAI NEI PARI MED**

“I belong [lit. “am”], together with my companions [sc. the other votive objects] to Trivia, the good (one) of the gods. Do not take possession of me.”

Vine (1998), while accepting **NE PARI MED** and **KOM MEOIS SOKIOIS**, reinterprets the remainder of the inscription significantly. His analysis of several of the forms of the text has a profound effect on the interpretation of the total inscription. Whereas Cristofani, and Mancini after him, see the inscription as a dedication to Trivia (**TRIVOIAD** or **TRIVOIAT**), an epithet of Diana and Marica, in whose sanctuary it was found, Vine convincingly refutes the Diana/Marica/Trivia equation on linguistic and orthographic grounds. His interpretation follows:

A. **AHUIDIES**

“Audiis/Audeius”

B. **ESOM KOM MEOIS SOKIOIS TRIBOS AUDEOM DUO[M] NEI PARI MED**

“I am, together with my three companions [the bowl/possession] of the two Audii (i.e., “I, together with my three companions, belong to the two Audii”). Do not take possession of me”.

The form **AHUIDIES** (Part A.) is most likely a nominative singular (it could be plural, though this appears less likely), marking the simple citation of the owner’s family name. It is not a Latin form, but is rather in some Oscan-

Umbrian (Sabellian) language of the Italic group. **AHUIDIES** has syncopated equivalents in Oscan as **ahvdiu** and **avdiis**, which correspond to Lat. *Audius* and *Audeius*, as Mancini notes (1997: 11 et seq.). Part B. of the inscription contains a number of interesting features, among which is the form **ESOM** (= CLat. *sum* "I am"), to which we return below. The form **KOM** equals CLat. *cum* "with"; **MEIOS SOKIOIS** (= *meīs sociīs* "(with) my companions") shows an older ablative plural ending *-ois*, previously found only twice in the inscriptional record, once in the Castor and Pollux dedication (*CIL* 1²: 4.2833): **CASTOREI•PODLOVQVEIQVE QVROIS** "To Castor and Pollux sons (of Zeus)"; and once in the highly controversial and debatable form **RIVOIS** "streams" (?) of the Duenos inscription (*CIL* 1².4); **TRIḐOS** (= *tribus* "three"), possibly **TRIḐOS**, modifies **MEIOS SOKIOIS**. Far from being the first part of the name *Trivia*, this is in fact part of the constituent headed by **SOKIOIS**; Vine's elegant analysis of **AUDEOM DUO[M]** matches the name **AUD(E)IUS** of Parts A. and B., thereby eliminating the need to justify the unexplainable *o* in Cristofani's (**TRIVOIAD**) and Mancini's (**TRIVOIAI**) readings of the text as containing the name for the goddess *Trivia*. As interpreted by Vine, **AUDEOM DUO[M]** shows the expected OLat. genitive plural form *-om* (later *-um*), which preceded the CLat. form *-orum* (cf. **DUUMVIRUM** "of a pair of magistrates"). This places the Garigliano Bowl text with other "talking inscriptions" (Agostiniani 1982), which uniformly express belonging with (*ego*)*sum* plus the genitive, not the dative as Mancini's reading requires; **NEI PARI MED** contains an older form of the negative particle *nē*, a form which is found elsewhere in the records of the archaic language, with diphthong preserved; **PARI** appears to be an imperative of *pariō* "I acquire", a 'third conjugation' verb (inf. *parere*), with frequent 'fourth conjugation' forms; **MED** is the common older accusative for "me", CLat. *mē* "me", and is a regular feature in archaic Latin inscriptions. The sense here of "do not take possession of me" for **NEI PARI MED** parallels almost exactly the injunction found in other talking inscriptions, most notably Duenos, where the bowl urges "do not let a bad man steal me" according to the interpretation of Rix (1985).

This brings us back now to **ESOM**, the form of Lat. *sum* "I am" which appears on Garigliano. The presence of **ESOM** in this inscription is of great potential significance for Italic and Indo-European scholarship. Few individual forms in the history of our discipline have attracted as much attention as have *sum* and its paradigmatic companions, *es*, *est*, *sumus*, *estis*, and *sunt*. Garigliano **ESOM** is the first specifically Latin example (even if Garigliano is non-Roman Latin) of another form of *sum*, and its appearance on this bowl could be useful in resolving the puzzle of the origins of this verb, not only in Latin and Italic, but also in PIE.

We can frame the problem specifically within the Italic context: given the confirmed presence of *esum/esú* in Oscan-Umbrian dialects,¹ and now *esom* in the (non-Roman) Latin of Garigliano, are we in a position to assert that **esom* can be projected back to Proto-Italic? Furthermore, does Garigliano *ESOM* help us to resolve the debate over the form of PIE "to be", specifically the first person singular, as either **h₁ésmi* (the standard reconstructed form) or **sem/som* (Meillet 1931; Bonfante 1932; Bader 1976; Schmalstieg 1980, to appear). It would appear that the Italic data are pointing in both directions at the same time, namely that the *sum* form suggests **sem/som*, while the *esom* form supports **h₁ésmi*, as do most of the IE languages, e.g., Skt. *ásmi*, Av. *ahmi*, Gk. *εἰμί*, Goth. *im*, Alb. *jam*, Hitt. *ešmi*, OIr. *am*, Arm. *em*, etc. Can these two forms be reconciled?

For the first question, does Garigliano *ESOM* confirm the Proto-Italic **esom*, the answer would appear to be that it does. Garigliano *ESOM* is roughly contemporaneous with the Oscan-Umbrian equivalents of this form. If it were a late analogical creation from a base form *sum*, it would have to have been formed by the same process independently in both the Latin-Faliscan and the Oscan-Umbrian branches of Italic, a most unlikely scenario.² Nor can it be considered a loan, first because, as Joseph & Wallace (1987) have noted, it antedates the period of close contact between the Latin-speakers and the speakers of the Oscan-Umbrian dialects. Second, a verb form like "I am" is an unlikely candidate for borrowing, even if considerable bilingualism and close contact between Latin and Oscan-Umbrian speakers were the norm. The pattern of Latin/Oscan-Umbrian loans is one of individual cultural items, not members of high-frequency paradigms.

Meillet, Bonfante, Bader, and most recently Schmalstieg have argued that Lat. *sum* and Oscan-Umbrian *sum/súm* are archaisms reflecting the IE state of affairs, and that *esum/esú/esom* are the result of various analogical processes.³ According to Schmalstieg (to appear), Latin *esom* is an analogical creation based on 2nd and 3rd person *es*, *est*. Those dialects in which the analogical form was created just didn't survive. By this argument (supported by Schmalstieg with comparative data from the Slavic first person singular forms), the analogical force of *es* and *est* was strong enough to generate the form *esom*, but not strong enough to keep it in the standard dialect. But in light of the evidence from the rest of the IE languages that the PIE form of the first singular was **h₁ésmi*, from which Proto-Italic **esom* can be uncontroversially derived, the postulation of a base form **sem/som* in PIE, and the creation of an analogical form *esom/esum* in Proto-Italic seems a needless complication. Given the evidence of Garigliano *ESOM*, there is now even more reason to believe that **esom* is a legitimate reconstruction for Proto-Italic, and that *sum* and its Oscan-Umbrian counterparts

are secondary, Schmalstieg's well-designed analogy notwithstanding. The sheer age of **ESOM** and its Oscan-Umbrian relatives (starting from the 6th C. BCE) point directly to ***esom**. I agree, then, with Joseph & Wallace, who argue, based on the four Oscan-Umbrian forms **esum** cited in their article, that "it does not appear feasible to treat **esum** as an analogical innovation".

Joseph & Wallace (1987) and Wallace & Joseph (1989) propose that the two forms of the verb "to be" in Latin and Oscan-Umbrian are the result of enclisis, a well-documented phenomenon in the history of the Italic languages: cf. Oscan **terminatust** "terminata est", **prúftúset** "posita sunt", Lat. *factust* "factum est", *moriundust* "moriundum est" or *scriptust* "scriptum est", and so on. They claim that an inherited base form ***esom** existed in two variants, one accented, one enclitic (687).

From a Proto-Italic form ***esom** (reconstructed on the strength of the Oscan-Umbrian forms **esum/esú** and now Garigliano **esom**), there are the following outcomes, partially adapted from Joseph & Wallace's (1987: 687) chart:

A. Latin	enclitic forms	accented forms
Proto-Italic	* ǵ-esom	* ésom
syncope	* ǵ-som	esom
raising of * o ⁴	* ǵ-sum	esom

There is then a generalization in Latin of the enclitic form *sum* over the accented form *ésom*.

B. Oscan-Umbrian	enclitic forms	accented forms
Proto-Italic	* ǵ-esom	* ésom
(→ South Campanian outcome)	esum/esú	esum/esú
syncope	sum/súm	esum/esú

There is then a generalization in Oscan-Umbrian of the syncopated enclitic form **sum/súm** over the accented forms **esúm/esú**.

Joseph & Wallace's chart can now be refined by the removal of the * in the Latin accented forms. Garigliano **ESOM** confirms the hypothesized Proto-Italic form ***esom**, and strengthens the case that Latin *sum* and Oscan-Umbrian **sum/súm** are secondary developments based on syncope under enclisis. Given the lack of further evidence among the IE languages for a form ***sem/som**, the case for PIE ***h₁ésmi** appears to have been made, illustrating once again (cf. Flobert 1991) the critical role which can be played by inscriptions in our understanding of preliterary languages.

Acknowledgments

I am grateful to Pierluigi Cuzzolin (Pavia), who brought the Garigliano Bowl to my attention, and to Brent Vine (UCLA), who commented on an early draft of this paper, and who is blameless for whatever faults and errors remain.

Author's address

Philip Baldi
Department of Classics & Ancient Mediterranean Studies
Pennsylvania State University
University Park, PA 16802, USA
phb@psu.edu

Notes

1. The forms *esum* and *esú* are attested in a number of inscriptions. The *esum* variant is found in the South Picene inscription of Campovolano of the 6th C. BCE (Marinetti 1985: 213–214, TE.4); the inscription of Vico Equense of the mid 6th C. BCE (Cristofani 1994: 381–384); the inscription of Nocera Alfaterna of the mid 6th C. BCE (Cristofani 1994: 381–384); and the inscription of Ager Signinus of the 5th to 4th C. BCE (Colonna 1994: 298–301). Three additional examples, from Capena, are spelled *esú*, the most recent one of which is discussed by Colonna (1989/90: 462–464), the other two of which appear already in Vetter (1953: 327, no. 356c and 1953: 329, no. 359d).
2. It is frequently noted in discussions of Lat. *sum* and related forms that Varro (*L.L.* 9, 100) claimed that *esum* was an earlier form of *sum*: *de infectis sum quod nunc dicitur olim dicebatur esum*: “Concerning the imperfective [sc. forms of the verb], that which is now pronounced *sum* previously was pronounced *esum*”. Until the recent appearance of *esum/esú/esom* in various Italic inscriptions, Varro’s words were usually discounted. Cf. Kent’s (1938) note to the Varronian claim in his Loeb translation (p. 519): “This form seems to have been invented by Varro to suit his argument: all the evidence is against its ever having existed”.
3. Schmalstieg’s arguments are the most specifically elaborated. While Bader (1976) relies on misinterpreted Tocharian data, and Meillet (1931) relies on a methodological principle (the famous dictum that it is isolated and exceptional forms in a language system that hold the keys to its history), Schmalstieg (to appear), in general agreement with Bonfante (1932), argues that inscriptional *esom* is an analogical innovation based on the 2nd and 3rd persons *es* and *est*. Even Schmalstieg, however, allows for the possibility of an Early Indo-European form **es-óm* hypothesized on the basis of internal reconstruction.
4. The evidence of the Garigliano bowl suggests that either *ESOM* is an exception to the rule whereby **o > u*, as are *longus* ‘long’ and *onus* ‘burden’; or else *ESOM* antedates the rule (cf. Early Latin *mol(c)ta > Lat. multa* ‘fine, levy’).

References

- Agostiniani, Luciano. 1982. *Le "iscrizioni parlanti" dell'Italia antica*. Florence: Olschki.
- Bader, Françoise. 1976. "Le présent du verbe 'être' en indoeuropéen". *Bulletin de la Société de Linguistique de Paris* 71.27–111.
- Birnbaum, Henrik & Jaan Puhvel, eds. 1966. *Ancient Indo-European Dialects*. Berkeley & Los Angeles: University of California Press.
- Bonfante, Giuliano. 1932. "Lat. *sum*, *es*, *est*, etc". *Bulletin de la Société de Linguistique de Paris* 33.111–129.
- Campanile, Enrico. 1985. "Riflessioni sui più antichi testi epigrafici latini". *AIÖN* 7. 89–99.
- Colonna, Giovanni. 1989/90. "Nuovi graffiti capenati". *Rivista di Epigrafia Latina, Studi Etruschi* 56.462–464.
- . 1994. "Ager Signinus (Lazio)". *Rivista di Epigrafia Latina, Studi Etruschi* 60.298–301.
- Cristofani, Mauro. 1994. "Sulle più antiche iscrizioni italiane della Campania". *La Preziosa Etrusca Nella Campania Meridionale* (= *Atti Delle Giornate di Studio Salerno-Pontecagnano, 16–18 Novembre 1990*). Istituto Nazionale di Studi Etruschi e Italici, Biblioteca di "Studi Etruschi" No. 28, 379–386. Florence: Olschki.
- . 1996. *Due testi dell'Italia preromana*. (Part I, *Per Regna Maricae*, pp. 7–32). Quaderni di Archeologica Etrusco-Italica, 25. Rome: Consiglio Nazionale delle Ricerche.
- De Simone, Carlo. 1980. "L'aspetto linguistico". Stibbe et al. 1990.71–94.
- Flobert, Pierre. 1991. "L'apport des inscriptions archaïques à notre connaissance du latin pré littéraire". *Latomus* 50.521–543.
- Joseph, Brian D. & Rex E. Wallace. 1987. "Latin *sum*, Oscan *súm*, *sim*, *esum*". *American Journal of Philology* 108.675–693.
- Kent, Roland G. 1938. Translation of M. T. Varro's *De Lingua Latina*. Loeb Classical Library. Cambridge: Harvard University Press.
- Lejeune, Michel. 1989. "Notes de linguistique italique. Notes sur la dédicace de Satricum" (XXXVIII); "Génitifs en *-osio* et génitifs en *-i*". (XXXIX). *Révue des Études Latines* 67.60–77.
- Mancini, Marco. 1997. *Osservazioni sulla nuova epigrafe del Garigliano*. (Biblioteca di Ricerche Linguistiche e Filologiche 42.1). Rome: Istituto de Studi Romani, Università degli Studi della Tuscia, Editrice "Il Calamo".
- Marinetti, Anna. 1985. *Le iscrizioni sudpicene* (I, *Testi*). Florence: Olschki.
- Meillet, Antoine. 1931. "Caractère secondaire du type thématique indo-européen". *Bulletin de la Société de Linguistique de Paris* 32.194–203.
- Prosdocimi, Aldo. 1985. "Sull'iscrizione di Satricum". *Giornale Italiano di Filologia* 15:2.183–230.
- Rix, Helmut. 1985. "Das letzte Wort der Duenos-Inschrift". *Münchener Studien zur Sprachwissenschaft* 46.193–220.

- Schmalstieg, William R. 1980. *Indo-European Linguistics: A new synthesis*. University Park: Pennsylvania State University Press.
- . To appear. "Latin *sum* and the Slavic first person singular present". *Historische Sprachforschung*.
- Silvestri, Domenico. 1993. "I più antichi documenti epigrafici del latino". *Campanile* 1993.97–118.
- Stibbe, Conrad M. et al. 1980. *Lapis Satricanus*. (Archeologische Studien van het Nederlands Instituut te Rome; Scripta Minora V). The Hague: Staatsdrukkerij (K. F. Treebus).
- Vetter, Emil. 1953. *Handbuch der italischen Dialekte*. Heidelberg: Winter.
- Vine, Brent. 1998. "Remarks on the Archaic Latin 'Garigliano Bowl' Inscription". *Zeitschrift für Papyrologie und Epigraphik* 121.257–262.
- Wachter, Rudolf. 1987. *Altlateinische Inschriften*. Bern & Frankfurt: Peter Lang.
- Wallace, Rex E. & Brian Joseph. 1989. "*Sum*: further thoughts". *Classical Philology* 84.319–321.
- Watkins, Calvert. 1966. "Italo-Celtic revisited". *Birnbaum & Puhvel* 1966.29–50.
- . 1989. "New Parameters in Historical Linguistics, Philology, and Culture History". *Language* 65.783–799.

CHAPTER 33

Comparative, Structural and Sociolinguistic Analyses of the History of the Romance Languages

Roger Wright
University of Liverpool

1. Historical linguistics and historical reality

The Historical-Comparative method of reconstructing preliterate languages proved so valuable that scholars in more recent areas of Historical Linguistics were tempted to use it too. But this has led to problems. Even in the original field of application it has to be admitted that the value of the method lay essentially in the lack of any competition; linguists, as linguists, could do nothing else but extrapolate backwards from later attested and potentially cognate languages. When the linguists in the tradition have been moved to speculate on the identity and whereabouts of the real people who spoke these reconstructed forms, however, they bump into the fields of expertise inhabited by archaeologists, and increasingly by such unforeseen collaborators as geneticists, and the results are not always such as to confirm the tentative conclusions reached on the basis of comparative historical linguistics alone. Thus even in these rarified realms, historical fact, however established, needs to be taken into account when we study historical linguistics; but whereas historical linguistics has for many years been taken to be part of linguistics, needing to take into account the advances and hypotheses made by linguistics in general, it has proved harder to establish the fact that it is also part of the study of history. And in one particular field, the study of the Early Romance spoken before the separate emergence of different Romance languages in different written forms, it seems to me that the

Historical Comparative method has outlived its usefulness.¹

Structuralism, however, the fashion that immediately followed the Historical Comparative, and which was exploited to help understand the data that the earlier method had discovered, deliberately ignores historical realities to an even greater extent. Indeed, it cannot be used to find out what happened, only (at best) to explain it once we know. Synchronic linguistic reality is there perceived to lie solely in internal structural oppositions. In synchronic phonetics, for example, even if a bilingual speaker pronounces the [ð] of English *lather* and Spanish *lado* identically (giving rise to identical acoustic spectrograms), the structuralist is forced to override this reality and claim they are not the same, since in English [ð] and [d] represent contrasting phonemes (as in *lather* versus *ladder*) but in Spanish they are allophones of the same phoneme in complementary distribution ([*lado*] is impossible); in synchronic semantics, two British English-speakers who refer to an identical vehicle with the word *bike* will be said by the structuralist to mean different things if one sees *bike* as incompatible with *motorbike* and the other sees it as superordinate to both *motorbike* and *pushbike*.

The point is that in this way neither the tradition of comparative reconstruction nor that of structuralism have seen the relationship between language and the real world as being essential primary data, which is probably why history itself was edged out of so much historical linguistics. Both the reconstruction of 'Proto'-Romance, as exemplified for present purposes by the work of Robert Hall Jr. and Robert De Dardel, and the present widespread delight in structural explanations of linguistic changes in Romance, seem to have been at least partly inspired by dissatisfaction with traditional, amateurish, explanations of change in terms either of historical events or of bilingualism and influences from other languages. And, indeed, both have clearly helped to raise the level of the discussion. In the first category, nobody would now seriously claim, as they used to be tempted to, that the development of the Romance languages was directly caused by the fragmentation of the Roman Empire, or that the change in the French accentual system was related to the Black Death; or even that the specific developments of (e.g.) French were caused by (e.g.) the Germanic spoken by the Franks. Substratum and adstratum languages can provide vocabulary, although even this is not in itself predictable, for neologisms are as often coined through derivational morphology and semantic change as through inter-lingual borrowing; but on other linguistic levels it is now no longer tenable, for example, to attribute to Oscan the developments that happened later in the area where Oscan was once spoken, and still less those that occurred in the Aragonese town of Huesca (< OSCA), as was once seriously suggested. The influence of Basque on Ibero-Romance is now agreed by all serious commentators (such as Echenique and

Trask) to have been almost negligible, for the languages are still, as they always have been, very different; the influence of Arabic on the development of Spanish, although it seems chill and unromantic to say so, is almost equally undetectable. Thus it was that explanations in terms of external historical events fell into disrepute, and explanations in terms of phonetic laws, structural pressures, typological drift, syntactic reanalyses and abductions, etc., largely came to take their place.

2. The value of historical sociolinguistics

But this decontextualized scenario has also slowly come in turn to seem partial, and the pendulum is swinging back; not to the ingenuous explanations favored in the past, but to the more sophisticated versions of historical sociolinguistics that have been growing up in the last few years, apparently to the surprise of some of our structuralist colleagues. The key development here has been the distinction made by James Milroy between innovation and change; an initial linguistic innovation may well have structural motivation, but the spread of that innovation throughout a community, which is what leads to a 'change', needs to be seen in historical and sociolinguistic terms. The sociolinguistics of bilingualism may have been deservedly relegated as an explanatory force in itself, but monolingual sociolinguistics remains more important than the structural tradition wants to give it credit for. And it is worth noting that bilingualism can still be used seriously to explain a change, even if it can no longer explain the original innovation. For example, it is no longer possible, for various reasons, to claim that Basque directly caused the Castilian development of initial [f-] to [h-] in words such as Latin FACERE, > Old Castilian [hadzer] (spelt *fazer*), which has since become Modern Castilian [aθer] (spelt *hacer*). On the other hand, all attempts at a solely structural or other internal explanation are going to fail to deal with the question of why this development happened in some geographical areas but not in others. The present state of play in this case (well expressed in Lloyd 1987: 267-73) is to accept that there was pre-existing linguistic variation among the allophones of Latin /f/, two of which were probably [f] and [ϕ], and that Basque-Latin bilinguals had a greater than average tendency, attributable to phonotactic features of their native Basque, to choose the more aspirated (less occluded) variants; but that this skewing of the allophone statistics could not have happened without the pre-existing allophony. In the absence of such already current variation within Latin for Basque to influence statistically, many other remarkable features of Basque have failed to leave any echo in the Romance of the

bilingual areas. Thus the initial innovation of allophonic variation for /f/ in Latin needs an internal explanation — if it was internally inexplicable, it would not have happened at all — and the satisfactory explanation of the subsequent spread needs knowledge of the world on the part of the historical linguist, that is, in this case, that the phenomenon happened particularly in the mouths of native speakers in the area of the Iberian Peninsula which provided the speech varieties that — for historical reasons alone — came to be treated as “standard” Spanish. A particularly interesting aspect of the sociolinguistic effects of bilingualism on language change in the modern world is unearthed in the painstaking analyses by Silva Corvalán of the Spanish of Los Angeles; developments within angeleno Spanish can be attributed to the nature of local Spanish-English bilingualism without particularly being attributable to English itself.

Another aspect of Romance development which has been greatly illuminated by historically-inclined sociolinguists in the recent past concerns the formation of what Trudgill calls an ‘interdialect’, and Romance specialists have tended to call a ‘koiné’, out of cohabiting mutually comprehensible dialects. This perspective has proved most useful so far in increasing our understanding of what happened to Spanish in the Americas. For many years it was normal to claim that Latin American Spanish descended from the speech of Andalucía, and as a generalization this was not unreasonable; Latin-American Spanish is more like that of Andalucía than that of any other separate region of the Iberian Peninsula. But there are features of andaluz that have not travelled to the New World: most obviously ‘ceceo’, in which standard /s/ and standard /θ/ are both [θ], and which in fact covers a greater extension of Andalusian geography than the ‘seseo’ which is normal in America, in which both are [s] (see the map in Narbona & Morillo-Velarde 1987: 60); and there are also features of other areas that have indeed crossed the ocean, such as the Asturian initial [ɲ-] found in parts of Argentina and elsewhere. The establishment of the American Spanish koiné has recently been traced by many experts on both sides of the ocean (e.g., in Lüdtke 1994), and is now largely understood, at least in broad outlines; they seem not to realize it, but this is essentially parallel to the work carried out by Trudgill on the English of the Southern Hemisphere. In each case the resulting speech, developed out of the separate but mutually intelligible dialects that contributed to its formation, particularly in colonial settlements and imperial armies, is a distinctive and essentially simplified combination of the highest common factors of the features of the contributory dialects. Penny (1987) has illuminated the reasons for the changes in the internal history of Castilian between the ninth and sixteenth centuries along these lines, and I have pointed out (Wright 1996) that the development of Early Romance outside Italy during the Roman Empire can also

be largely explained in terms of the formation of an interdialect of this type among the areas outside the Italian Peninsula, a koiné that eventually came to be influential in Italy itself also; and this helps to explain why the history of the later developments of the Roman Empire seem to imply that, as well as being a time of considerable linguistic evolution, it was a time of convergence rather than divergence.

3. Unfortunate Romance corollaries of the comparative method

At this point the application of modern language-internal sociolinguistics, employed to explain the sociohistorical trajectories of successful innovations and the formation of koinés, crashes head on into one of the central beliefs of the comparative method as applied to Romance; the belief, enshrined in the image of the tree diagram as a model for the evolution of separate languages from a common source, that fragmentation is both geographically discrete and historically linear. The specialists in Proto-Romance reconstruction have modelled their theoretical activity on Proto-Indo-European; in the PIE field it makes some sense to provide tree diagrams of divergence, both because the time scale is enormous and because the speakers of Proto-Indo-European can legitimately be seen to have spread out as distinct communities, often separated from each other by considerable distances in the sparsely inhabited world of the prehistoric centuries, and lacking any kind of unifying force in a common written standard. None of this applies in the Early Romance world, within which there was continuous travel and intercommunication, and a unifying standard conveniently available in written form to act as a centripetal magnet for all styleshifting. But the force of the analogy has proved so strong that the language-internal variation which is known to be an inevitable feature of all wide speech communities has been completely ignored within this tradition (or at best dismissed as a 'kitchen-sink' theory), and there has instead arisen a tendency for every conceivable reconstructed variant to be allotted to a discrete language variety localized on the soil on which the historical descendants of the variant concerned are known to survive in the attested medieval languages or in the present day. Thus it is that a very early separation from 'Proto'-Romance in general for '(Proto)-Sardinian', dated — if at all — to three centuries before Christ, is seen as gospel in this tradition, and the main bifurcations between geographically separable Romances are dated to well before the end of the Roman Empire. The only evidence for this is the analogy with PIE, and a feeling that it must be so. Historical evidence, on the other hand, overwhelmingly suggests that there was a monolingual (if

variable and complex) community for almost a millennium after that.

I have examined elsewhere (Wright 1995: chapter 5) another baleful consequence of the analogy with PIE (there are probably others unrecognized): the inappropriate use of the asterisk. The asterisk is a device used sensibly within the reconstruction of PIE (popularized by Schleicher, although, as Koerner has shown, not invented by him) to indicate the reconstructed existence of a word, presented in its hypothesized phonetic form. In discussing a preliterate age, this device makes sense. Within Romance Historical Linguistics, however, it has become normal to use the asterisk with unattested forms of otherwise well attested lexical items (e.g., *COMPERARE, between Latin *comparare* and Italian *comprare*). In such a case what has been reconstructed is a pronunciation, but what is unattested is a written form that would be isomorphic with that reconstructed pronunciation if Latin spelling were a phonetic script, but of course Latin spelling was not a phonetic script and the lexical item concerned is indeed often copiously attested. Over 90% of the asterisks used (in the 23 studies investigated) turn out to be in this category, added to an unattested written form of a lexical item which is attested in some other written form; this approach is by definition not applicable to the PIE centuries, in which the question of what a word's written form would have been if the speakers had been literate is irrelevant in the extreme. The practice of adding asterisks to 'Proto'-Romance forms presented in putatively isomorphic orthography has often led to confusion between a lexical item (as a vocabulary unit), its orthographical shape (on parchment), its phonological shape (in the mind), and its phonetic shape (vibrations in the air), and thus to many inconsequential and dangerously misleading arguments that would never have arisen without the unfortunate overapplication of this aspect of the comparative method to a literate community.

4. The discovery of Latin pronunciation

The origins of the comparative reconstruction of Proto-Romance are entirely honorable. They lay originally in the recognition of the fact that traditional Latin orthography was not a phonetic transcription of its writers' vernacular; this fact is still occasionally capable of eluding Latinists even to this day, as they, for example, describe the orthography of a Late Latin text under the heading of 'Phonetics'. Robert Hall, Jr., in particular, realized, correctly, that the texts of the first millennium A.D. could be no reliable guide as to the vernacular pronunciation of the period. So he and his colleagues used the attested facts of mid-medieval Romance languages, on the partly-justifiable assumption that at least in

the twelfth and thirteenth centuries these were fairly close to being a phonetic transcription of the habits of their scribes, as the prime data for reconstructing backwards and establishing the nature of the common phonetics (or phonology) of the language from which all the Romance languages had sprung. This was valuable and important, and Robert Hall's great achievement. What he and his colleagues had done was to discover the pronunciation of Latin, and for this they ought to deserve eternal gratitude; but the problem was that they did not know that this was what they had done. Instead they said they had discovered the phonetics (or phonology) of 'Proto-Romance', conceived of as being a different language from the Latin of the Empire; for they assumed, without argument or even consideration, that the Latin of the Empire was at that time pronounced in the manner of later medieval Latin (that is, a specified sound per written letter, as Latin is indeed now pronounced, but as has almost certainly only been the case since the invention of medieval Latin as a separate entity in the early ninth century A.D.). This assumption is still not dead: see for example Cull (1995).

5. The myth of a Proto-Romance 'system'

Following this line of thought, the reconstructors of Proto-Romance deduced that since all texts were Latin, and thus in a different language, they were therefore irrelevant to the study of Proto-Romance. De Dardel's latest book (1996) takes this conclusion to remarkable lengths; for De Dardel, texts of the time are 'parole', whereas he is interested in the Proto-Romance 'langue'. This attitude is to some extent acceptable if we are studying phonetics and phonology alone, for traditional orthography was the only one taught (and probably often taught logographically, one word — or abbreviation — at a time); although even then it is advisable to have our proposed chronology of phonetic developments in rough alliance with the order and existence of attested orthographical peculiarities in the texts. It is partly acceptable in the study of nominal morphology, for we know from the surviving manuals on *Orthographia* that the inflections of nouns and adjectives were taught during the process of learning to write; but such deliberate evasion of the textual evidence is not acceptable if we are studying verbal morphology, syntax, semantics, or even vocabulary, which were less open to conscious pedagogical control than the orthography and much of the nominal morphology were. It is the function of the grammarian down the ages to confuse the evidence, certainly, but writers are not going to change features of their own speech of which they are not even aware, such as the semantic changes that may have occurred recently in the words they are using, or the pragmatic value of the

different available methods of expressing passive meaning. An appropriate appreciation of the relationship between speech and writing is a central crucial aspect of all historical analysis of literate languages, for the relationship can vary greatly between different social and historical contexts, and the analyst needs to be aware of the relationship in the historical context of the language being analyzed. This has indeed been studied for the Early Romance-speaking world (by Banniard, Herman, Varvaro, myself, and several others), but is discarded as irrelevant within the Romance comparative tradition. De Dardel, indeed, invents his own examples to exemplify the syntactic features he claims to have 'reconstructed'. In particular, he claimed (in an earlier article, De Dardel 1989) to have reconstructed four separate changes of basic word order in Romance between the third century B.C. and the sixth century A.D., including one which is not otherwise attested in Europe (he regards this not as a drawback of the theory but as its crowning achievement). Actual texts do not attest these developments, but that is seen as neither here nor there, even though it is exceptionally difficult for writers to vary their natural word order in their own compositions, at least in the absence of explicit pedagogical instruction (and the idea that Latin should have verbs at the end of its sentences, for example, is not one which is stressed in the pedagogical grammars of this age, which largely refrain from mentioning syntax at all). As Bob Blake has shown on several occasions, ostensibly Latin (or 'Latinate') texts of central medieval Spain, in the period preceding the orthographical reforms, show Romance word order patterns.

The structuralist idea that each language is a 'system' seems to be to blame for all this unnecessary confusion. Languages are in reality mostly composed of accidental conglomerations of contingentially concomitant details, only some of which are at times organized architecturally into inter-relating partial substructures — indeed, it may be debatable whether any language other than a genuine isolate is in reality an isolatable entity at all — but an intellectual obsession with finding intrinsic language-wide systematicity can obscure this unwelcome fact. In this case, the one simple mistaken idea, that the reconstruction of 'Proto'-Romance phonology was that of a separate language different from Latin (rather than just being what it was, the pronunciation of Latin), is a misapprehension easy to dispel with a simple knowledge of basic historical and social facts about the Late Empire and the Early Middle Ages (which are not at all "Dark" Ages: we actually know a surprising amount about this period, each century of which saw the writing of more texts than survive from each century of the Empire). Even so, it seems to have been what led to the further idea that there were two whole distinct languages during the Roman Empire, not just the two hypothesized separate phonologies. Therefore, in the inexorable logic of De Dardel

(1996), there must have been a separate syntax too, since all autonomous languages are said in the structuralist approach to be complete language systems, with their own separate syntactic structure; and since, in his view, by definition Proto-Romance was not Latin, therefore the syntax shown in texts cannot by definition have been the same as that of speech, and therefore improbable conjectures, justified by the use of rigorous comparative procedures alone, have greater validity than the direct textual evidence which was so thoughtfully provided for us by those who supposedly spoke in this reconstructed manner. This is all achieved in the name of the dignity of comparative method. Yet let us look at the method as employed elsewhere: comparative historical linguistics in other fields has achieved wonderful progress, of course, but not by refusing to consider the available textual evidence. New data that turn up in the sands of Turkey are avidly examined and exhaustively analyzed, not ignored. Specialists in Proto-Bantu would love to have a seventh-century text; specialists in Proto-Romance reconstruction have available for study a large number of seventh-century texts, particularly from the Iberian Peninsula, and prefer to ignore them.

6. Why Romance fragmentation happened when it did

For these reasons it seems fair to conclude that historical comparative method in the Romance languages has probably achieved all that it ever will. We should, of course, forever honor the discovery of Latin pronunciation, and of otherwise unattested vocabulary and derivational morphology. The Romance languages and structural linguistics can still co-exist: for they provide a wonderful testing ground for all the potentially explanatory internal structural proposals that come to be made within historical linguistics as a whole, although it is still the case that Spanish and Italian, in particular, tend to be used in the most patronizing and even insulting manner as 'test cases' for improbable theories developed on the basis of English. But if we first look at the Romance historical data pre-theoretically, we are likely to conclude that the most interesting feature of the development of Romance is probably not internal at all, and thus intrinsically not amenable to a merely structural explanation. For contrary to an apparently widespread assumption made by historical comparativists, language splits are not inevitable; and yet — at a much later date than that usually proposed, but still undeniably occurring — we have here a well documented real-life scenario in which what was originally one language has come to be thought of as being several different languages, even in an area which — unlike the vast geographical extension of the original Indo-European languages — continued literate, and

whose communities continued in direct personal contact, throughout (with the conceivable exception of Early Romanian-Romance speakers, since we do not really know where they were). It now seems reasonable to claim that the fragmentation can only be explained through a knowledge of history allied to a knowledge of what is plausible in sociolinguistic terms. The reconstruction theory can offer branching trees, but no convincing dating for the branches (and glottochronology and lexicostatistics are at best unconvincing). The theory is effectively unable to help with the time dimension. Structuralism, on the other hand, is also stumped by geography. Structural analyses are often able to explain the initial rise of an innovation, but not why it should have been adopted in some places but not in others, a variability in practice which is the defining characteristic of the fragmentation at the focus of our lens; typological analyses are especially impotent in such cases, since if typological explanations of change, even when vaguely rephrased as examples of 'drift', are taken to be predictive, then there is no way to explain the undoubted fact that the Romance languages are not all the same as each other (nor are all the Germanic languages, nor all the Indo-European languages, for that matter); and if areal quasi-universals are used to explain the overruling of supposedly inexorable typological developments (as in Campbell 1980), we are in effect conceding thereby that social and historical events can explain crucial aspects of linguistic evolution.

In the particular case of the conceptual separation of the Romance languages from Latin and from each other, a knowledge of external history, including conscious language planning, leads us to a scenario which is more straightforward and likely. More, that is, than either the quasi-mystical belief in perpetual bilingualism (sometimes rephrased as diglossia) and very early geographical fragmentation proposed by the application of historical comparative method, or the reliance on decontextualized typological or other forces proposed within the structural tradition. The Latin-speaking world, also known as the Early Romance world, continued for many centuries to develop, quite rapidly on phonetic, morphological, and lexical levels, quite slowly on a syntactic level — as is usual — but until at least the sixth century A.D. there were centripetal cultural forces leading to convergence which could counteract any evolutionary potential for divergence. This chronology in itself lessens the possibility of substratum influence; there may well still have been features attributable to Iberian in the Latin of the third-century Peninsula, or features attributable to Phoenician in the remarkable third-century data that has been recently emerging from Bu Njem in Tunisia (see e.g., Adams 1994), but these would be far less likely to have been still present five centuries later (as Arabic influence was much more apparent in 13th century Spanish than in that of the 18th C.). This monolingual Early

Romance world, containing variation of a wide but normal language-internal kind such as happens in any large community, probably continued at least to the end of the eighth century, with contemporary written texts still largely intelligible when read aloud (even if older ones may not have been: cf. Herman 1996). The subsequent separation of Latin from Romance is due to a historically contingent occurrence: that is, it was a consequence of the establishment of an artificially revived Latin based on the elementary grammar of Donatus, the more advanced linguistic analyses of Priscian (only recently brought West from Byzantium), and the Anglo-Saxon habits of spelling-pronunciation (only recently brought East over the channel). This newly normativized Latin began to be established around 800 A.D., in Charlemagne's Empire, although it took a long time to be general practice all over the Romance world, and it was this ambitious language-planning exercise which introduced the Latin-Romance distinction which has seemed natural for the last few centuries. In due course, and once again for contextually explicable reasons that can be historically traced, different Romance-speaking areas began to write their vernaculars in different ways (that is, with different *scripta*, as has come to be said), following experiments initially carried out in Northern France which probably had a limited aim at first, being intended to aid Germanic speakers who knew the new Latin to read or sing in a vernacular manner coherently or in unison. Eventually, for political reasons connected with growing mid-medieval nationalism, every Romance kingdom, around the year 1200, seems to have wanted to establish its own written form as a symbol of national identity and pride. This nationalism helped finally hypostatize the separate Romance languages. The process is not over yet, for the same train of thought is visible in the autonomías of Spain in the 1990s, several of whom would like, for political reasons, to establish their own local spelling norms and thus claim to have a separate language ('Asturian', 'Murcian', etc.). The point being made here is that all these developments were, from a solely linguistic point of view, quite unnecessary and in any sense unpredictable; the English, French, and Chinese experiences show that it is possible for a conceptually monolingual state of affairs to continue despite huge linguistic variability, perhaps even indefinitely, for the variations that can be held within such elastic monolingualism can be enormous, both socially and geographically, without such conceptual fragmentation necessarily occurring; the Romance developments are easy to explain if we can see the historical background, and hard or even impossible to understand if we cannot.

7. Why this conclusion is significant

We can say that historical comparative method has achieved valuable results in the Romance field, but there is no need to push the method further any more; that structural and typological approaches have obvious value in synchronic analyses, and in several Pan-Romance diachronic developments, but their overdetermining nature means that they are at a loss to explain one of the most salient features of Romance development, which is the mid-medieval conceptual fragmentation; and that it is time now for all workers in this field to accept that sympathetic analyses of the texts and their writers, of sociolinguistics, and indeed of simple linguistic history, all have an important role to play in our understanding of the evolution of Romance languages. That these external details happen to be available to the Romance specialist to a greater extent than to students of any other linguistic family is not something for which the Romance specialists need to apologize; we just need to exploit it in a sensible way, and help specialists in other families to accept that the particularities of a long-past historical context might, if only they knew what it was, have explained developments that seem inexplicable on comparativist and structural grounds alone.

Author's address

Roger Wright
Department of Hispanic Studies
University of Liverpool
Liverpool L69 3BX, UNITED KINGDOM
rhpwri@liv.ac.uk

Notes

1. This paper brings together and develops several inchoate ideas scattered in earlier studies, particularly Wright 1982, 1995, 1996 (interdialects, and convergence), and 1997 (the splits of the 12th and 13th centuries).

References

- Adams, James N. 1994. "Latin and Punic in Contact? The case of the Bu Njem Ostraca". *Journal of Roman Studies* 84.87-112.
- Banniard, Michel. 1992. *Viva Voce: communication écrite et communication orale du IVe au IXe siècle en Occident latin*. Paris: Institut des Études Augustiniennes.
- Blake, Robert. 1991/96. "Syntactic Aspects of Latinate Texts of the Early Middle Ages". *Latin and the Romance Languages in the Early Middle Ages* ed. by Roger Wright, 219-232. London & New York: Routledge (paperback reprint, University Park, Pennsylvania: Pennsylvania State University Press).
- Campbell, Lyle. 1980. "Explaining Universals and their Exceptions". *Papers from the 4th International Conference on Historical Linguistics* ed. by Elizabeth C. Traugott et al., 17-26. Amsterdam & Philadelphia: John Benjamins.
- Cull, Naomi. 1995. "Reconstruction of the Proto-Romance Syllable". *Historical Linguistics 1993: Selected Papers from the 11th International Conference on Historical Linguistics* ed. by Henning Andersen, 117-132. Amsterdam & Philadelphia: John Benjamins.
- De Dardel, Robert. 1989. "L'hypothèse d'une base OVS en protoroman". *Probus* 1.121-143.
- . 1996. *A la recherche du protoroman*. Tübingen: Max Niemeyer.
- Echenique Elizondo, María Teresa. 1987. *Historia lingüística vasco-románica*. 2nd ed. Madrid: Paraninfo.
- Herman, József. 1990. *Du latin aux langues romanes*. Tübingen: Max Niemeyer.
- Koerner, Konrad. 1989. *Practicing Linguistic Historiography*. Amsterdam & Philadelphia: John Benjamins.
- Lloyd, Paul M. 1987. *From Latin to Spanish: Vol I*. Philadelphia: American Philosophical Society.
- Lüdtke, Jens, ed. 1994. *El español de América en el siglo XVI*. Frankfurt: Vervuert (= *Bibliotheca Ibero-Americana*, 48).
- Milroy, James. 1993. *Linguistic Variation and Change: On the historical sociolinguistics of English*. Oxford: Blackwell.
- Narbona Jiménez, Antonio & Ramón Morillo-Velarde Pérez. 1987. *Las hablas andaluzas*. Córdoba: Cajasur.
- Penny, Ralph. 1987. *Patterns of Language Change in Spain*. London: Westfield College.
- Silva Corvalán, Carmen. 1994. *Language Contact and Change: Spanish in Los Angeles*. Oxford: Oxford University Press.
- Trask, R. Larry. 1996. *The History of Basque*. London & New York: Routledge.
- Trudgill, Peter. 1986. *Dialects in Contact*. Oxford: Blackwell.
- Varvaro, Alberto. 1991/96. "Latin and Romance: Fragmentation or restructuring?" *Latin and the Romance Languages in the Early Middle Ages* ed. by Roger Wright, 44-51. London & New York: Routledge (paperback reprint, University Park, Pennsylvania: Pennsylvania State University Press).

- Wright, Roger. 1982. *Late Latin and Early Romance in Spain and Carolingian France*. Liverpool: Francis Cairns.
- . 1995. *Early Ibero-Romance: Twenty-one studies on language and texts from the Iberian peninsula between the Roman Empire and the thirteenth century*. Newark, Delaware: Juan de la Cuesta Monographs (Estudios Lingüísticos no.5).
- . 1996. "Latin in Spain: Early Ibero-Romance". *The Origins and Development of Emigrant Languages* ed. by Hans F. Nielsen & Lene Schoesler, 277-98. Odense: Odense University Press.
- . 1997. "Linguistic Standardization in the Middle Ages in the Iberian Peninsula: Advantages and disadvantages". *De mot en mot: Aspects of medieval linguistics* ed. by David A. Trotter & Stewart Gregory, 267-81. Cardiff: University of Wales Press.

CHAPTER 34

Romance Historical Morphology and Empty Affixes

Martin Maiden
University of Oxford

Analysis of word-structure into 'morphemes', component signs, notoriously runs aground on 'meaningless' elements of morphological structure. Thus while Spanish *humareda* "cloud of smoke" may be analyzable into stem *hum-* (cf. *humo* "smoke") and derivational suffix *-eda*, the residual *-ar-* is meaningless. 'Empty morphs' are among various form-meaning asymmetries prompting abandonment (cf. Matthews 1972; Anderson 1992; Beard 1995) of assumptions that words are concatenations of signs (morphemes) in favor of a view that only 'lexemes', not bound grammatical morphemes, are signs, and that apparent structure within words results from phonological operations on stems quite separate from semantic processes. Recent years have also seen rising interest in 'morphomic' structures: that is, in properties of morphological structure viewed as occupying an autonomous level mediating between morphosyntax and phonological realization of lexemes. An example of a morphomic structure is Aronoff's analysis (1994: 36–39) of the 'third stem' in Latin verbs, a distinctive stem-form about which no phonological generalization is possible (it varies unpredictably from verb to verb), yet which, whatever its form, systematically and consistently occurs in certain parts of the inflectional and derivational paradigm of the verb over which no semantic generalization is available. This is not the place to attempt anything approaching a detailed exposition and assessment of these arguments; my aim, rather, is the modest one of suggesting that the work of various Romance historical linguists of the first seventy-five years of this century deserves the serious attention of those interested in issues of 'autonomous' morphology, because the analyses the Romanists invoke seem to require the existence of what are in modern terms 'morphomic' affixal constituents, constituted by non-null and non-discontinuous elements of morphological structure which have no independent meaning.

Romanists were examining 'empty' affixal elements already in the 19th C.,¹

but I shall begin with a line of investigation commenced in the 1900s by Menéndez Pidal who, on several occasions over some fifty years (notably 1905; 1926: 337–44; 1953) studied semantically empty ‘unstressed suffixes’ in Ibero-Romance (and beyond). I cannot review his accounts of their origins (a task performed, with major fresh insights, by Craddock 1969, Malkiel 1972) but we may say briefly that various Ibero-Romance words (especially in north-western dialects) display an ‘unstressed suffix’, devoid of concrete meaning, immediately following their stem. The ‘unstressed suffix’ consists (Menéndez Pidal 1905: 387–397) of a vowel (usually [a], also, more rarely, [e], [i] or [o]) followed by a consonant (in descending order of frequency, [r], [n], [g], [l], [m], [d], [b]). Examples are Spanish *cáscara*, ‘tanning bark, grape skins’, *alicántara* ‘mythical venomous snake’, *bonítalo* ‘bonito’, coexisting with *casca*, *alicante*, *bonito* of the same meaning; *ciénaga* (Salamanca *ciénago*) ‘marsh’, *relámpago* ‘lightning flash’, *pezpítalo* ‘wagtail’, *luciérnaga* ‘firefly’, from older *cieno*, *relampo*, *pezpita*, Latin LUCERNA ‘lamp’; *galápago* ‘tile mould’, probably related to *galapo* ‘block for twisting rope’; New Castilian *támara* ‘threshed straw’, with synonymous *tamo*; Asturian *áscuara* ‘ember’, *pómpara* ‘bubble’ (compare Spanish *ascua*, *pompa*), *pítara* ‘whistle’ (from the stem of *pitar* ‘to whistle’); Salamancan *tállaro* ‘stalk’ (alongside *tallo*). Menéndez Pidal’s thesis is that the ‘unstressed suffixes’ originate in words which happen to end in the (semantically empty) sequence ‘unstressed vowel + consonant + vowel’, such as *ángaro* ‘beacon’, *búcaro* ‘fragrant clay’, *jácara* ‘merry ballad’, *nácara* ‘kettle drum’, *rábano* ‘radish’, *orégano* ‘marjoram’, *plátano* ‘plane tree’,² *búfalo* ‘buffalo’, *sábalo* ‘shad’, *piélagos* ‘ocean’, *espárrago* ‘asparagus’, etc. There was no existing motivation for analyzing these as ‘stem + suffix’, since there were no ‘stems’ *ang-*, etc., and no (unstressed) ‘suffixes’ *-aro*, etc.; Menéndez Pidal underscores (1905: 399f.) the meaninglessness of these ‘suffixes’, observing that they often constitute a mere ‘morphological adornment’ (‘adorno morfológico’), but only in his 1926 treatment does he distinguish terminologically between ‘adornments’ and ‘meaningful’ suffixes, calling the former ‘suffixed increments’.

The question of the ‘meaninglessness’ of unstressed suffixes proved problematic. In 1905 Menéndez Pidal observed that they sometimes had vague popular or rustic nuances but in 1953 he adopts the new position that their source was a suffix existing in a pre-Latin language and originally possessing a ‘collective’ meaning vestigially continued into Romance. But on this point Malkiel observes (1972: 319; see also Craddock 1969: 85) that: ‘the author seems unaware of the fact that he is referring to ever-changing, iridescent CONNOTATIONS and that, so far as straight DENOTATION is concerned, the elements at issue

simply lack any that could have been captured in fifty years of intensive study — a circumstance which might well have encouraged him to question, if not downright deny their suffixal rank.” And indeed it would be extremely difficult to identify any common factor even in the ‘connotations’ reported. It is tempting to say that ‘unstressed suffixes’ are, precisely, ‘increments’, or ‘adornments’, empty morphological vehicles capable of bearing whatever connotations one might invest them with (cf. also Prati 1942: 85 for similar facts in Italian), rather than items that can be listed in association with a well-defined meaning. Malkiel adds a further twist to the question of the possible ‘meaning’ of the unstressed suffixes by arguing (1972: 323–330) that a source for the diminutive value sometimes associated with *-alo* was survival into popular Romance speech of an unstressed diminutive **-olo* from Latin *-ULU(M)* (cf. also Menéndez Pidal 1953: 50). By Malkiel’s account, the [o] of *-olo* was supplanted by [a], a vowel both widely represented in the unstressed penultimate syllable of many words of Greek origin, and especially resistant to the syncope which tended to affect Ibero-Romance intertonic vowels. Replacement of original intertonic vowels by [a]³ also helps explain the emergence of other words in *-ago*, *-aro*, *-ano*. But any association between ‘diminutiveness’ and *-alo* seems rapidly to have been abandoned, and subsequent generalization of *-alo* as an ‘unstressed suffix’ goes hand in hand with that of ‘meaningless’ *-ago*, *-aro*, *-ano*, etc. Whatever semantic generalizations might be made about the earliest stages of the suffixes, what seems clear is that their subsequent generalization transcends any identifiable semantic motivation. Of the ‘unstressed suffixes’, which Malkiel had relabelled ‘nominal augments’, he writes:

The semantic weakness, bordering on emptiness, of the nominal augments places these increments at the very periphery of affixation, in the vicinity of suffixoids,⁴ interfixes (the semantic deadwood par excellence), and integrating or serializing (‘*einreihend*’) suffixes, such as *-al* of Sp. *roble* *dal* ‘grove of oak trees’ and, at least initially, of E. *histor-ic-al* (originally a synonym of *histor-ic*). Like the suffixoids, the augments are not always neatly detachable from the ‘primitives’; in fact, the words into which they enter seldom qualify for the rôle of ‘derivatives’, but rather represent re-shaped, stylized variants of the older formations — variants exhibiting special elaboration or sharpening of the word-final segment (1972: 334).

By likening the ‘nominal augments/unstressed suffixes’ to ‘suffixoids’, ‘interfixes’, etc., Malkiel implies that they are to be viewed as elements of *morphological* structure. Yet they seem to fulfill a primarily phonological function, for he observes (1972: 331, 334) that the spread of the nominal augments was probably promoted by a fashionable, playful, liking for proparoxytonic stress

patterns, and the unstressed suffixes provided a means of doing this. It is striking also that Menéndez Pidal assigns 'unstressed suffixes' to morphology, but at the same time (1953: 55) says that they have 'simply auditory effect'. Other commentators effectively consider them a matter of phonology: Hanssen (1913: §384) has reservations about their 'suffixal' nature, alluding simply to 'terminaciones esdrújulas' ('proparoxytonic endings'), while Michaëlis (1905) dubs them an 'acrecento eufónico' ('euphonic augment'). Modern 'separationist' models (cf. Anderson 1992: 68f.) would have little difficulty in describing such phenomena synchronically as adjunction of phonological material to stems. Indeed, one may wonder about the basis for assigning such phenomena to 'morphology'.

Although this question was not directly addressed by scholars who discussed 'unstressed suffixes', it seems to me implicit in much of their work that if the 'suffixes' were not morphemes in the classical sense, some notion of 'suffix' as a structural element of words remained desirable in order to give an account of certain formal, phonological, properties of those 'suffixes', and in particular their striking paradigmatic coherence. The elements *-aro*, *-ano*, *-alo*, etc., are interesting in this respect, for they, or rather their consonants, readily substitute each other both within the same speech variety and across geolinguistic space (cf. also the intersubstitutability of suffixal vowels described in Prati 1942: 82). Menéndez Pidal shows that, alongside *pifaro* "fife", from German *Pfeifer*, there evolved *pífano*; Old Spanish had *búfano* corresponding to *búfalo* "buffalo"; Portuguese *cángaro* "crab" corresponds to Asturian *cánganu*; and there is evidence that *zángano* "idler, drone" once also occurred as *zángalo* and *zángaro*. These cases do not represent a mere array of alternative strategies all serving the common function of creating proparoxytones, for there happen to be no paroxytones (at least in the relevant senses) ***pifo*, ***bufo* etc.; rather, we see a paradigmatic analogical influence between the suffixes which has nothing to do with whatever stress patterns they may confer. Menéndez Pidal gives many similar examples (extending his purview into Italo-Romance), ascribing the intersubstitutability (1953: 55) of these elements partly to their acoustic resemblance (they comprise [a] + nasal or liquid consonant + vowel), and partly to their shared semantic 'emptiness'.⁵ Still such an analysis might favor a purely phonological approach: the 'unstressed suffixes' are meaningless and have characteristics describable phonologically — so why invoke 'morphology' at all? But if the unstressed suffixes are just phonological, they occupy a peculiar place within Spanish phonology: they do not arise from any historical phonological process, apparently involve a unique⁶ pattern of intersubstitutability between intervocalic [n], [l], and [r], and constitute a rather peculiar phonological 'constituent' (the nucleus, but

not the onset, of the penultimate syllable plus the whole of the final syllable) which is otherwise unparalleled — unless, that is, we consider the characteristic shapes of *morphological* elements: the majority of derivational suffixes have precisely the form ‘nucleus + final syllable’ (cf. *delgadeza* “thinness” [del-ga-ðe-θa] = stem [delgað] + suffix [eθa]; *gatito* “little cat” = [ga-ti-to] = stem [gat] + suffix [i-to], etc.). The behavior of unstressed suffixes as described by Menéndez Pidal and others would seem to make better sense if we said that certain formal properties generally characterize suffixes, independently of any meaning that may be associated with them.

The existence of a class of empty but ‘suffix-like’ elements is also strongly implicit in Meyer-Lübke’s account of French historical word-formation. He observes (1921 = 1966: 7) that one cannot speak of ‘suffixes’ in cases where either the ‘stem-word’ does not exist, or the final syllable lacks definite meaning, citing recurrent word-final elements such as *-eil(le)* (from Latin diminutive -ICULU(M)/A(M)) in French *soleil* “sun”, *pareil* “similar”, *orteil* “toe”, *corbeille* “pannier”, *corneille* “crow”, etc., where *-eil(le)* has no distinct function, and there exist no corresponding “stem-words” **sol-*, **par-*, **ort-*, etc., ‘in the relevant meaning. Recurrent word-final elements like *-eil(le)*, labelled ‘Ausgänge’, are distinct from true ‘suffixes’, yet seem to be more than accidentally recurrent rhyming word-ends, for they are substitutable (1921 = 1966: 10f.) by suffixes proper (i.e., by ‘suffixes’ conceived as meaning-bearing morphemes). Under ‘Suffixwechsel’ or ‘Suffixverdrängung’ (= ‘suffix exchange’, ‘suffix displacement’) Meyer-Lübke details substitutions involving *-ier* (from Latin adjectival/agentive -ARIU(M)) for earlier *-er*, in *bachelier* “(knight-) bachelor”, *bouclier* “buckler, shield”, *collier* “necklace”, *écolier* “schoolboy”, *pilier* “pillar”, *sanglier* “boar”, *singulier* “singular”, *soulier* “shoe”, etc. He treats these cases as substitutions of *suffixes* (rather than of *Ausgänge*), and adduces a partial semantic motivation in that *-er* characterizes substantivized infinitives and is thereby (supposedly) associated with ‘abstractness’, while *-ier* words are more ‘concrete’. Appeal to supposed ‘abstract’ associations seems merely an attempt to elicit some semantic motivation for the change, and to assign it thereby to ordinary ‘analogy’ involving form-meaning matchings: but the very fact that there already existed a longish series of concrete words in *-er* might equally have disfavored association with abstractness, and Meyer-Lübke’s account is not compelling. In fact, it may simply be the numerical superiority of *-ier* over *-er* forms, abetted by their partial phonological similarity, which favored the substitution. What is striking about these examples is that whilst some had transparent internal structure (e.g., *école* “school” + suffix *-er*) others (e.g., *piler*, *sangler*, *souler*, *bachelor*) did not, so that by Meyer-Lübke’s own criteria they did not comprise

a 'stem-word' + 'suffix', but an initial element plus Ausgang. That an Ausgang can be substitutable by a 'true suffix' suggests that, at some level, Ausgang and suffix are identical: in other words, *souler*, etc. must manifest an internal structure equivalent to 'stem + suffix', even though 'stem' and 'suffix', taken separately, are contentless. Even where such substitutions have clearer semantic motivation, they still imply 'empty' morphological analysis of the word whose ending is substituted. Replacement of Ausgang *-eille* in *ouaille* "flock of sheep" by the 'collective' suffix *-aille* (Meyer-Lübke 1921 = 1966: 12) yields *ouaille* despite the non-existence of stem ***ou-*. 'Collectivity' inheres in the entire lexeme *ouaille*,⁷ and this presumably motivated replacement of *-eille* by *-aille*, but for such replacement to happen, two things are necessary: that existing collectives in *-aille* (e.g., *mangeaille* "victuals") be analyzed as bimorphemic, permitting the 'detachability' of *-aille* as collective suffix, and that a similar 'bimorphemic' — better, 'bimorphic' — analysis ('stem' + Ausgang) be assigned to *ouaille*, allowing substitution of *-eille* by *-aille*. Other analogical changes presupposing assignment of form independently of meaning are also described by Meyer-Lübke 1921 = 1966: 12f.; 25: replacement of the final element of *cisoir* "scissor" by *-eau* (*ciseau*), supposedly under the influence of *marteau* "hammer", despite the lack of stems *cis-* or *mart-*, and of any particular association between the very common ending *-eau* and "tools"; appearance of etymologically unexpected *-on* in *taon* "horsefly", attributed to resemblance with a group of semantically related words in *-on* such as *frelon* "hornet", *bourdon* "bumblebee", and other insect-names, despite the absence of stems of appropriate meaning in *ta-*, *frel-*, *bourd-*; internally opaque *saumon* "salmon", apparently interpreted as *saum* + *on*, spawning *saumier* "salmon gaff"; Old French *oiseus* "bird" hatching derivatives *oisance*, *oiserie*, etc., despite the non-existence of independent stem ***ois-*.⁸ A similar example (1921 = 1966: 141), this time involving a prefix, is the substitution in Old French of *bes-* "twofold" for *ob-*, so that *oblong* "oblong" yields *beslong*. Here, although *ob-* had no independent value, *long* was an independent form meaning "long"; but no such internal structure was available in *oblif* "oblique" < Latin OBLIQUU(M), which nonetheless also yielded *beslif*.

It is Malkiel who initiates (cf. Gsell 1981: 2), within Romance linguistics, a concern with 'empty morphs' as a major theoretical issue. For those 'empty' elements which separate lexical stems from derivational suffixes in Ibero-Romance, Malkiel devises the term 'interfixes'. His 1958 article asserts the *autonomous* status of 'interfixes' as contentless morphemes, against the dogma that morphemes necessarily have well defined functions, but his argument for recognition of 'interfixes' relies principally on criteria of descriptive economy.

Possible treatment of the interfix + suffix sequence as a single, unanalyzable, suffix (so that Spanish *humareda* might comprise *hum-* + suffix *-areda*, rather than *hum-* + interfix *ar* + suffix *-eda*) is rejected because the result would greatly expand the inventory of suffixes. Malkiel argues (1958: 184f.) that it would be more practical to relax the rigid definition of the morpheme, applying the term not only to the minimal meaningful unit, but also to empty portions of words remaining after analysis into 'morphemes', reserving the term 'residual' or 'marginal' morpheme for these 'subordinate' and 'exceptional' linguistic elements.⁹ Malkiel's study transcends procedural elegance and economy of synchronic structural analysis by showing that the 'interfixes' are, in effect, 'psychologically real' for speakers, and that their existence conditions morphological change. An example (1958: 142) is *jabato* "young boar", formed from *jabalí* "boar", in which *-al-* was historically an integral part of the stem (< Arabic *ġabal*) and for which there was no other evidence for a stem *jab-*.¹⁰ Malkiel writes that this "extreme case of regression", which presupposes an "ingenuous" analysis by speakers (cf. *egu-ar* alongside *egu-al-ar*, both "to equal"), gives "a certain air of reality" to the interfix *-al-*, showing that it represents something more than an artefact of formal analysis. Among other examples indicative of diachronic reanalyses involving postulation of empty structural elements are (1958: 117, 124, 134): *peinetilla* now interpreted in Mexico as a diminutive of *peine* "comb" (i.e., *pein-* + interfix *et* + diminutive *-illa*) as well as of *peineta* "ornamental comb"; the coexistence in Portuguese and Spanish of *tosse/tos* "cough" with unrelated *tósego* or *tósigo* "poison", yielding Old Spanish and Portuguese *tossegoso* "having a cough", where the semantically vacuous *-eg-* arose from a false analysis of *tósego* as *tos* "cough" + interfix *-eg-* + derivational suffix.

Such examples of speakers' capacity to discern 'empty structure' are striking, but need not be essentially different from phonological 'abductions', where some aspect of actually 'underlying' phonological structure is interpreted as 'derived'. These examples are not, then, necessarily inconsistent with affixation viewed as a phonological operation on stems. But Malkiel's writings are also rich in examples which seem to invoke a 'paradigmatic dimension' of affixation, in which the history of 'empty' affixes seems explicable only by recognizing a distinct class of 'affixal' structural elements capable of exercising *analogical* influences on each other. He shows (1958: 140) that from *testera* "top of head" was derived *testerada* "blow with the head" which further yielded *testarada*. To account for *testarada* Malkiel argues that *-er-* was first reanalyzed as an empty 'interfixal' element attached to *test-*, then supplanted by the more frequently occurring, and equally empty, interfix *-ar-*. To assert that *-er-* and *-ar-* represent

phonological operations on a stem *test-* would provide no way of capturing what *-er-* and *-ar-* have in common and what permits replacement of one by the other,¹¹ namely their 'affixhood'. Similar intersubstitutability between interfixes of different etymological origin, *-eg-*, *-ag-* (and sometimes *-ig-*), is described in Malkiel (1949: 152–54, 180).

A more dramatic example of what amounts to analogical change in the (near) absence of meaning is Malkiel's account (1966) of *-ido*. Briefly, Old Spanish had a system of conjugation-dependent past participial affixes following the stem, comprising one of three vowels + [d]: *-ado* (1st conjg. *cantar cantado*), *-udo* (2nd conjg. *atrever atrevudo*), *-ido* (3rd conjg. *pedir pedido*). It also acquired a (temporarily) flourishing series of novel forms in *-ido* with loose connotations of 'weakness' or 'inadequacy', and not directly derived from verbs: e.g., *desperido* (< Latin *DE/EXPERITUS*) "exhausted", *adolorido* "doleful", *descolorido* "colorless", *desabrido* "tasteless", *desertido* "uninhabited", *atelerido* "shivering", *desleído* "weak", *escaldrido* "sly". Malkiel writes (1966: 335) that

The three elements — call them allomorphs if you wish — *-ado*, *-udo*, *-ido* formed a neat, well-integrated triad in the conjugational paradigm of medieval Spanish, a triad for which I propose the term 'first string'. The point is that *-ado* and *-udo* (initially, to the exclusion of *-ido*) also formed part of a smaller second string — denominative formations presupposing no intermediate or cognate verb. [Among them: *-ado* indicating resemblance, color or relation (*aindiado*, *ahazendado*); *-udo* indicating excess, abundance, plethora (*barbudo*, *peludo*, *pacienzudo*)] As a result of this inequality, Spanish had to its credit a longer first string of deverbal *-do* suffixes and a shorter second string of denominative *-do* suffixes, from which *-ido* was noticeably absent. There developed, in its suffixal system, something reminiscent of the diachronic phonemicists' celebrated [...] 'case vide'. This hole could be most readily filled with the flotsam and jetsam of isolated words, many of them [...] of controversial origin; with moribund past participles in *-ido*, retreating before the new series in *-ecido* and qualifying, henceforth, only for limited adjectival uses; and with lexical innovations, especially if these could be effectively polarized vis à vis the *-udo* group [...].

In fact, what is described here (see also Malkiel 1974: 76) is a little different from the structuralist 'case vide': if, say, a language has a set of phonemic voice oppositions for consonants, except that it only has in the labiodental series /f/, the subsequent rise of /v/ can be ascribed to its slotting into a kind of structural 'grid' based on existing meaningful oppositions. The point about the *-ado*, *-udo*, *-ido* series in the verb is that the distinction between them is one of — inherently semantically arbitrary — conjugational class, a matter of pure form (cf. Aronoff 1994: 45f.), devoid of meaning — and Malkiel even proposes calling them

'allomorphs', a term reserved in the structuralist tradition for non-meaningful paradigmatic variations in the form of meaningful elements.¹² Yet this meaningless 'first string' supposedly creates the model for addition of *-ido* to the 'second string', a series of forms whose differences in meaning, whatever they may be — and Dworkin (1985) argues that Malkiel overstresses the role of expressive meaning — has nothing in common with conjugation. What is actually being postulated is a matter of pure morphological form, a kind of analogy in which a semantically vacuous paradigmatic array *-ado -udo -ido* favors creation of a new member, *-ido*, in the partially identical paradigmatic series already comprising *-ado* and *-udo*. The notion of semantically vacuous proportional analogies in paradigms is rather startling, in that such analogies usually involve matchings of form to meaning. If there is any 'meaning' here, it is internal to morphology: the fact that the series of suffixes *-ado* and *-udo* implies *-ido*.

The Old Spanish pronunciation of the suffixes *-az-*, *-iz-*, *-uz-* (e.g., *aguza* "sharpness"; *granizo* "hail"; *escaseza* "scarcity"; *hornazo* "Easter cake stuffed with hard boiled eggs"; *hilaza* "yarn"; *espinazo* "backbone"; *hinchazón* "swelling") with voiced *z* [dz], rather than etymologically expected voiceless *ç* [ts], leads Malkiel (1971; also 1981) to the hypothesis that the aberrant phonological development of the suffix reflects structural pressure from the other Spanish suffixes, virtually all of which happen to be characterized by a vowel followed by a *voiced* consonant (1971: 3f.): "The overwhelming majority of Hispano-Romance suffixes either leaned on a resonant as their pillar or axis [...] or contained such a consonant, or combination of consonants, as moved, in the course of normal sound changes, into the class of resonants or of voiced obstruents...": *-al*, *-il*, *-el*, *-uelo*, *-ar*, *-ero*, *-oro*, *-ura*, *-or*, *-ano*, *-in(o)*, *-ón*, *-uno*, *-ajo*, *-ijo*, *-aje*, *-ado*, *-edo*, *-ido*, *-udo*, *-iego*, *-ambre*, *-umbre*, (*‘-alo*, *‘-aro*, *‘-ago*). In fact only the 'rustic, coarse, comic' *-aco*, *-ieco*, *-ueco*, *-ico*, *-uco*, *-acho*, *-ete*, *-ito* contradict this generalization. Malkiel goes on to say that within this context it is permissible to argue that "by switching from *-ç-* to *-z-* in the cases of truly suffixal *-açar*, *-uçar*, *-aço*, *-iço*, *-açón*, and *-eça*, the speakers not only allowed the majority group within the total arsenal of suffixal tools available to them to contaminate this cluster of five sibilant-governed suffixes, but, by doing so, separated the sibilant-dominated group from a — semantically — very misleading phonic association with the residue of suffixes left to display some voiceless obstruent...". Whatever role avoidance of undesirable connotations played in replacement of [ts] by [dz], clearly — assuming the general correctness of Malkiel's analysis¹³ — the major determinant was structural pressure from the mass of suffixes, characterized as they were by voiced consonantal 'pillars'. They may be 'true suffixes' in that each has an identifiable meaning, but taken

together, as Malkiel's analysis presupposes, they have in common only that they are suffixes.¹⁴ And it is striking that a few, necessarily empty, 'suffixoids',¹⁵ (e.g., *erizo* "hedgehog"; *tomiza* "esparto rope"; *sazón* "sowing time" — in the absence of independent stems *er-*, *tom-*, *s-*) also undergo the change. In short, we have 'analogical levelling' whose determinant is not an iconic matching of one form to one meaning, but a matching of one form to one (suffixal) structure. It is extremely difficult to see how these 'paradigmatic' dimensions of suffixhood could ever be captured in an approach which views suffixation merely as a phonological operation on the stem, and does not fundamentally differentiate affixation from phonological operations like metathesis or revowelling. The diachronic facts suggest the existence of a 'list' of items which are indeed basically meaningless, but sensitive to the common property of 'suffixhood'.

This brief sketch seeks to do no more than suggest that the work of earlier Romance linguists may be worthy of attention by those interested in questions of autonomous morphology. In particular I cannot, within the space and chronological limits available, do justice to Malkiel's contribution, and I have not discussed here Malkiel (1974), which postulates semantically arbitrary "deep morphology", abstract templates to which morphological changes conform.¹⁶ The analyses I have reviewed above actually lend considerable support to the view that morphological structure needs to be considered in abstraction from any meaning which could be associated with 'morphemes' — in fact they make little sense if 'meaning' is invoked. But a notion of 'affix' as a component structural element of words, and the listing of forms which constitute affixes, seem desirable to account for the changes I have examined. A valid objection is, of course, that 'separationist' models are concerned with *synchronic*, not diachronic, morphology, and that the treatment of affixation as just another phonological operation might remain perfectly valid from a synchronic perspective. But the problems of accounting for the diachronic facts without recourse to some notion of autonomously morphological 'affixes' suggest that synchronic principles must also exist which make reference to such entities.¹⁷

Author's address

Martin Maiden
Trinity College
Oxford OX1 3BH, UNITED KINGDOM
martin.maiden@modern-languages.oxford.ac.uk

Notes

1. Malkiel (1958: 107–116) surveys earlier analyses of 'empty' affixes. Craddock (1969: Ch. 2) examines treatments of the subject from the 19th C. until the 1960s.
2. In some cases the *-an-* 'suffixes' first attracted other endings (cf. Malkiel 1972: 328). Thus *cuévano* 'hamper', *huérfano* 'orphan', etc., continue COPHINU(M), ORPHINU(M), etc.
3. Malkiel relates these facts to parallel developments in Tuscan.
4. 'Suffixoids' are equivalent to Bloomfield's (1933: 240) 'primary suffixes', for example the *-er* of 'hammer', 'rudder', which appears to be the same *-er* found in, say, 'leader', but whose 'stem' ('ham-', 'rud-') has no independent existence.
5. Menéndez Pidal talks, in fact, of a coherent group proper to early Mediterranean languages. But this 'coherence' is not merely etymological, for the suffixes appear to have progressed 'in lockstep', in their historical transmission (see Menéndez Pidal 1953: 36), presumably long after losing any original collective meaning.
6. Many of the examples Menéndez Pidal cites contain [r], [l], and [n] in the 'stem', which raises the possibility of a dissimilatory motivation, but it may be precisely the existence of alternative suffixal forms containing other consonants that makes the dissimilatory substitution possible — on this cf. Maiden (1997 and forthcoming).
7. Given that *ouelle* was, in principle, an unanalyzable 'stem', one might have expected *-aille* to be attached to that stem, yielding ***oueillaille*.
8. Some of these examples may represent the kind of 'back-formations' (cf. 'to sculpt', from 'sculptor') dismissed by Beard (e.g., Beard 1998: 56f.) as tending to be conscious, logical, operations, outside the domain of grammatical derivation and belonging to what he terms 'lexical stock expansion'. I doubt whether the production of, say, *saumier* from *saumon* was ever a conscious or logical operation, and it is far from clear that one can properly draw a line between conscious and grammatical operations. But the central point is that such examples show that speakers *do* attribute sign-like internal morphemic structure to words, and that such analysis has an important role in diachronic change. Compare also Aitchison (1994: 131).
9. In an intriguing remark (1958: 161f.) on the relation between origin and function of interfixes, Malkiel states that a grammatical element may come into being to fill a gap, smoothe a transition, or acclimatize a loan word. The notion of teleological genesis of interfixes (against, as Malkiel observes, much North American structuralist thinking) seems to me, however, insufficiently developed, and its full implications difficult to assess. The ensuing discussion (also 1958: 173–176) of 'antihatic' elements (e.g., *-t-* in *chapeau-t-er*; Salamanca *lengu-ar-udo*) implies a repertory of empty morphs available for phonological purposes such as hiatus avoidance (cf. also Maiden forthcoming). Malkiel also suggests (1958: 165–171) that interfixes may be exploited to avoid homonymy.
10. So far as I can ascertain, the postulated 'stem' never appeared in unaffixed form (as ***jabo*), so that creation of *jabato* seems to involve direct replacement of 'affixal' material to the right of *jab-* by *-ato*.
11. Of course their partial phonological similarity plays some role in determining the substitution, and one cannot exclude the possibility of a sporadic and purely phonetic change. But there is no general phonological principle in Spanish changing *-er-* into *-ar-*, and by the general premises of Malkiel's argument we have to say that this substitution concerns the *affixes* *-er-* and *-ar-*.

12. Craddock (1969: 55) discerns in Munthe (1905: 323) the implication that *-aru* and *-anu* are 'variants of the same morpheme'.
13. Here I merely explore the implications of Malkiel's analysis, assuming its general correctness. I intend elsewhere to explore an alternative hypothesis whose implications are equally interesting, namely that voicing is the result of a sporadic sound change which somehow 'saw' suffixal structure. It is a commonplace that sound change can be sporadic and restricted to certain *lexemes*, but a change affecting a *suffix* (regardless of the lexical item in which that suffix occurred), would suggest that suffixes have an autonomous status like that of independent *lexemes*.
14. Malkiel's theoretical interpretation (1970) of Canellada's account (1944) of Cabraniego hypothesizes that affixal productivity is a function of the conformity of the affix to a type of phonological structure characterized by a consonantal "pillar" and one of a series of paradigmatically alternating consonants. It hardly seems possible to identify an overarching common meaning for the class of affixes displaying such structure. But original and theoretically significant as Malkiel's study is, it contains much that is purely hypothetical and is arguably not based on the most reliable of sources (cf. Gsell 1981: 3).
15. 'Suffixoids' (for which see Malkiel 1966; 1970: 74f.) correspond closely to Meyer-Lübke's 'Ausgänge' and the term suggests that things that 'look like' suffixes in some respects *are* 'suffixes'. Compare also Malkiel's observation (1966: 324) that "the growth of suffix and suffixoid go hand in hand", citing for example *Cluniacus* > Osp. *grunniego* "Cluniac"; *PERSICA* > dial. *persiega* "peach"; *charnière* > *charniega* "hinge"; and *FANIQA* > Astur. *faniega* "bushel".
16. Unlike Malkiel, I think his Romance verb examples show paradigmatic differentiation of stems, rather than adjunction of "interfixes". Be this as it may, the fundamental insights of the 1974 study ought to become familiar to all interested in notions of 'autonomous' morphology. See further Maiden (1992, 1996).
17. There are parallels between many of the facts considered in this study and arguments from apparent 'truncation' processes as discussed by Aronoff (1976) or Corbin (1987). Aronoff's account is criticized by Anderson (1992: 280f.) who argues that an alternative purely phonological account is superior (also Aronoff 1994: 7). But in discussing French examples adduced by Corbin (1987), Anderson's appeal to 'phonology' seems at best vague:

...other instances of truncation seem clearly to involve morphological non-constituents [...];

- a. *virus* — *viral*; *cactus* — *cactée*; *rectum* — *rectal*; *tétanos* — *tétanique*
- b. *liquide* — *liquéfier* — *liqueur*; *stupide* — *stupifier* — *stupeur*
- c. *rigide* — *rigidifier* — *rigueur*; *humide* — *humidifier* — *humeur*
- d. *certain* — *certitude*; *caillou* — *caillasse*
- e. *charité* — *charitable*; *hérédité* — *héréditaire*; *vanité* — *vaniteux*
- f. *adroit* — *adresse*; *maladroit* — *maladresse*

In none of these cases have we any reason (apart from truncation) to believe that the truncated material is a morphological unit in French: truncation appears here to be applying to phonologically characterized parts of words. In the absence of clear instances of the situation illustrated by Aronoff, where only a morphological specification will specify the truncated material, we conclude that rules of truncation do not provide clear evidence for imposing morphological structure on words.

Now clearly the 'truncated' elements do not conform to traditional notions of the morpheme-as-sign, but to ascribe them to 'phonology' is to overlook the possibility of a contentless morphological structure. The phonological explanation involves postulating an ad hoc and disparate array of otherwise unmotivated phonological operations; about the only thing that (most of) these truncations have in common is that they appear to involve modifying phonological material to the right of a 'constituent' comprising the initial syllable plus the onset of the following syllable. To my knowledge no other rule of French phonology makes reference to such an unlikely 'constituent'. What this phonological analysis misses is that truncation seems to presuppose a pattern precisely characteristic of French morphological paradigms, namely that words can be analyzed into a 'stem' + 'ending', where the stem, if it ends in a consonant, is not coextensive with any syllable, for example (*nous*) *partons* [par-tō] comprising the stem *part-* + inflectional ending *-ons*, or *parler* [par-le] comprising the stem *parl-* + inflection *-er*. These facts are phonological only in so far as they concern the characteristic phonological shapes of *morphological* entities.

References

- Aitchison, Jean. 1994. *Words in the Mind*. Oxford: Blackwell.
- Anderson, Stephen. 1992. *A-Morphous Morphology*. Cambridge: Cambridge University Press.
- Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. Cambridge, Mass.: MIT Press.
- . 1994. *Morphology By Itself*. Cambridge, Mass.: MIT Press.
- Beard, Robert. 1995. *Lexeme Morpheme Base Morphology*. Albany, New York: SUNY Press.
- . 1998. "Derivation". *The Handbook of Morphology*, ed. by Andrew Spencer & Arnold Zwicky, 44–65. Oxford: Blackwell.
- Bloomfield, Leonard. 1933. *Language*. New York: Holt.
- Canellada, María José. 1944. *El bable de Cabranes*. Madrid: Revista de Filología Española supplement 31.
- Corbin, Danielle. 1987. *Morphologie dérivationnelle et structuration du lexique*. Tübingen: Niemeyer.
- Craddock, Jerry. 1969. *Latin Legacy versus Substratum Residue: The unstressed "derivational" suffixes in the Romance vernaculars of the Western Mediterranean*, University of California Publications in Linguistics, 53.
- Dworkin, Stephen. 1985. *Etymology and Derivational Morphology: The genesis of Old Spanish denominal adjectives in -ido* Tübingen: Niemeyer.
- Gsell, Otto. 1981. "Das sprachwissenschaftliche Œuvre Yakov Malkiels". *Ibero-Romania* 13.11–29.
- Hanssen, Friedrich. 1913. *Gramática histórica de la lengua Castellana*. Halle: Niemeyer.
- Maiden, Martin. 1992. "Irregularity as a Determinant of Morphological Change". *Journal of Linguistics* 28.285–312.

- . 1996. "The Romance Gerund and System-Dependent Naturalness in Morphology". *Transactions of the Philological Society* 94.167–201.
- . 1997. "La dissimilation à la lumière des pronoms clitiques en roman". *Zeitschrift für romanische Philologie* 113.531–562.
- . (forthcoming) "Phonological Dissimilation and Clitic Morphology in Italo-Romance". *Phonological Theory and the Italian Dialects* ed. by Lori Repetti.
- Malkiel Yakov. 1966. "Genetic Analysis of Word Formation". *Current Trends in Linguistics* 3 ed. by T. Sebeok, 305–64. Paris & The Hague: Mouton.
- . 1949. "Studies in the Hispanic Infix *-eg-*". *Language* 25.139–181.
- . 1958. "Los interfijos hispánicos". *Miscelánea Homenaje a André Martinet II* ed. by Diego Catalán, 109–99. Madrid: Gredos.
- . 1970. *Patterns of Derivational Affixation in the Cabraniego Dialect of East-Central Asturias*. Berkeley: University of California Press.
- . 1971. "Derivational Transparency as an Occasional Co-determinant of Sound Change. A new causal ingredient in the distribution of *-ç-* and *-z-* in ancient Hispano-Romance". *Romance Philology* 25.1–52.
- . 1972. "The Rise of the Nominal Augments in Romance. Graeco-Latin and Tuscan clues to the prehistory of Hispano-Romance". *Romance Philology* 26.306–34.
- . 1974. "New Problems in Romance Interfixation (I). The velar insert in the present tense (with an excursus on *-zerl -zir* verbs)". *Romance Philology* 27.304–55.
- . 1981. "A Hidden Morphological Factor Behind Instances of Erratic Distribution of *-ç-* and *-z-* in Old Spanish?". *Romance Philology* 35.105–29.
- Menéndez Pidal, Ramón. 1905. "Sufijos átonos en español". *Bausteine für eine romanische Philologie: Festgabe für A. Mussafia* ed. by Elise Richter, 386–400. Halle: Niemeyer.
- . 1926. *Orígenes del español*, Madrid: Suárez.
- . 1953. "Sufijos átonos en el Mediterráneo occidental". *Nueva revista de filología hispánica* 7.34–55.
- Meyer-Lübke, Wilhelm. 1921 [=1966]. *Historische Grammatik der französischen Sprache. II Wortbildungslehre*. Heidelberg: Winter.
- Michaëlis de Vasconcelos, Carolina. 1905. "Algumas palavras a respeito de púcaros de Portugal". *Bulletin Hispanique* 7.140–196.
- Munthe, Å. W. 1899. "Ein neuer Beitrag zur Kenntnis der asturischen Mundarten", *Zeitschrift für romanische Philologie* 23.321–25.
- Prati, Angelico. 1942. "Antisuffissi". *L'Italia dialettale* 18.75–166.

CHAPTER 35

Markedness and Morphosyntactic Change Revisited

The case of Romance past participle agreement

John Charles Smith
University of Oxford

1. Introduction

Andersen (1997) claims that “markedness relations can be observed in every variety of linguistic change, from its inception to its completion, both in the relations among variants and in the relations that define the plethora of categories that typically condition the gradual process by which newer forms replace older correspondents”. In this paper, I shall discuss the notion that the differential disappearance of agreement between the past participle and a direct object in the Romance compound past tenses formed with the auxiliary ‘have’ may be an instance of a linguistic change which is sensitive to markedness. I shall show that this view was anticipated a century ago by Bréal (1897), long before ‘markedness’ entered the vocabulary of linguistics, but will claim that the relatively unspecific nature of ‘markedness’ is the Achilles’ heel of this type of explanation for language change.

2. Markedness and morphosyntactic change

The idea that morphosyntactic change is actualized first in unmarked contexts is not a novel one. For instance, changes which affect main clauses before subordinate clauses, such as the change of word order from SOV to SVO in many Germanic languages, have long been seen as examples of the unmarked context being in the vanguard of change (for some recent discussion, see

Matsuda 1998). However, the notion was given particular prominence by Timberlake (1977), in his discussion of two diachronic processes. The first is the disappearance of a subject-to-object raising rule in Finnish complex sentences involving a participial complement clause, which leads to the underlying subject of the participial clause always appearing in the genitive as opposed to any of the cases normally associated with objects. The second is the replacement of the Russian objective adverbial genitive found with certain verbs by the accusative. The former change (1977:156) is "actualized earlier for constituents which are relatively more subjectlike" (pronouns before nouns; agentive or animate nouns before non-agentive or inanimate nouns). The latter change (1977:160) takes place according to a "hierarchy of individuation, the extent to which an object is considered as an individual" (proper nouns before common nouns; animate nouns before inanimate nouns; concrete nouns before abstract nouns; singular before plural). Timberlake (1977:168-169) concluded that the changes are "accomplished earlier in contexts that are unmarked for the innovation and later in contexts that are marked for the innovation". However, "the concept of markedness, or naturalness, must be understood with reference to the particular change involved". More recently, Andersen (1990) has provided further evidence for the view that morphosyntactic change is actualized first in unmarked contexts, in his examination of a development in Polish whereby forms of the auxiliary verb 'to be' become verbal endings. Once again, this development takes place earlier or more readily in contexts which can broadly be defined as 'unmarked' — not only morphological contexts (present tense before preterite, indicative mood before conditional mood, singular number before plural number, plural number before dual number, third person before other persons, first person before second person), and grounding contexts (main clauses before subordinate clauses, asyndetic clauses before syndetic clauses), but also genres (prose before poetry, expository prose before artistic prose, secular content before religious content), media (spoken language before written language), and styles (casual before formal) (see Andersen 1990: 10).

3. Markedness and the Romance compound past tenses

In Smith (1989), I identified five actualizations of the reanalysis of HABERE + past participle as a compound past tense form (the loss of agreement between past participle and direct object; the recasting of the relationship between the auxiliaries 'have' and 'be', often leading to the loss of the latter; the emergence of a fixed element order in which the auxiliary precedes the participle; the

decrease in the intercalation of items between auxiliary and participle; and the dissociation of auxiliary 'have' and the lexical verb signifying possession). I suggested, tentatively, that several of these actualizations tended to take place first of all in unmarked contexts. For instance, the extension of the auxiliary 'have' to unaccusative verbs which originally formed these tenses with 'be' may take place first of all when the subject follows the verb. According to Ronjat (1937: §583), in many Occitan dialects we find *siés tomba au sòu* 'you-are fallen to the ground', but *a tomba de nèu* 'has fallen of snow'. For the Limousin dialect, Roux (1895: 103) gives *A tomba una plueja rajousa* 'Has fallen a rain torrential' and *Una plueja rajousa es tombada* 'A rain torrential is fallen'; whilst Sicre (1909: 42), in his grammar of the Languedocien dialect of Foix, quotes the examples *A bengut qualqus* 'Has come someone' and *Qualqus es bengut* 'Someone is come'. Montes Giraldo (1976) notes the retention of 'be' as the auxiliary of some compound past subjunctive forms (but not indicative forms) in Colombian Spanish. In some Italian dialects of southern Lazio, *tenere* 'hold' has replaced *avere* 'have' as verb of possession in the present tense, but not in the preterite (Seifert 1935: 108-109). For fuller discussion of the sense in which these changes may be seen as sensitive to markedness, see Smith (1989: 320-321).

4. Markedness and object-participle agreement

4.1 Bréal's anticipation of the markedness hypothesis

However, perhaps the most interesting candidate for analysis in terms of markedness is the *règle de position* found in standard French and many French dialects, whereby the past participle agrees with a preceding direct object, but not with one which follows (for some discussion of this rule, see Smith 1993a). A similar pattern of agreement is found in some other Romance languages (see Smith 1995b). As early as 1897, Bréal (1897: 224-225) proposed that the *règle de position* of French was an example of this principle, although he did not use the term 'markedness'.¹ He notes (loc. cit.):

La contagion fournit, je crois, la véritable explication d'un fait de la langue française qui a beaucoup occupé nos grammairiens: le changement du participe passé passif en participe actif. Dans ces phrases: «J'ai reçu de mauvaises nouvelles, j'ai pris la route la plus directe», *reçu*, *pris*, ont aujourd'hui le sens actif, qu'ils doivent au voisinage de l'auxiliaire *avoir*. La preuve qu'ils ont le sens actif, c'est qu'en langage télégraphique je dirai: «Reçu de mauvaises nouvelles. — Pris la ligne directe.»

Là est, si je ne me trompe, la raison de cette règle de non-accord qui a donné lieu à tant d'explications embarrassées. La vérité est que le participe, par contagion, est devenu actif. Il fait corps avec son auxiliaire. Mais comme il a fallu du temps pour opérer ce changement, comme les anciens tours sont longs à se perdre, et comme la moindre dérogation au train ordinaire leur est un prétexte pour se maintenir, le changement en question ne s'est imposé qu'avec la construction la plus fréquente, celle que nous sommes habitués à considérer comme la construction normale. Partout ailleurs, la langue se montre fidèle à l'ancienne grammaire.

4.2 *Agreement hierarchies*

The *règle de position* is not the only sense in which the disappearance of object-participle agreement is differential. In Smith (1995b), I suggested that the differential disappearance of this type of agreement might be explained, at least in part, by sentence-processing strategies and the notion of recoverability, noting that the data enable us to establish a number of implicational hierarchies (in the sense of Greenberg 1963), of the form: "if, in a given language or dialect, the past participle agrees with a direct object of type X, then it will also agree with a direct object of type Y". These hierarchies are presented below (the notation $X > Y$ is to be interpreted to mean that agreement with X implies agreement with Y, but not necessarily vice versa).

a) Position of Direct Object:

Following > Preceding

b) Identity of Preceding Direct Object:

{	Topics Interrogatives Exclamatives	}	> Relatives > Clitic Pronouns
---	------------------------------------------	---	-------------------------------

c) Person of Clitic Pronoun:

{	First Person Second Person Third Person Reflexive	}	> Third Person Non-Reflexive
---	---------------------------------------------------------	---	------------------------------

d) Number and Gender of Third-Person Non-Reflexive Clitic Pronoun:

Masculine Plural > Feminine Plural > Feminine Singular

I also suggested that each of these hierarchies had a functional explanation, rooted in sentence-processing strategies, which can be summarized as follows (see Smith 1996; 1997: 1100-1101).

Hierarchy a). A direct object which precedes the verb is not in canonical or unmarked position, so that a sentence containing such an object will be more difficult to parse than one in which the direct object follows the verb. Agreement of the participle may therefore facilitate processing by serving to 'flag' both the immediately following empty object position and the number and gender of the item elsewhere in the sentence to which it should be linked.

Hierarchy b). In the case of Topics, Interrogatives, and Exclamatives, all the information required in order to determine the referent of the direct object is present in the same sentence as the participle. Similar arguments could be advanced in the case of relatives — they require an antecedent, which is normally found in the same matrix sentence. However, here the risk of ambiguity is greater, as there may be more than one plausible antecedent. Clitic pronouns, on the other hand, are much more difficult to process — they may be deictic, and so require pragmatic resolution, and can present problems even when endophoric (that is, anaphoric or cataphoric). Moreover, in the Romance languages, when a clitic ends in a vowel, this is often elided before the initial vowel of the auxiliary 'have', giving rise to identical surface forms for clitics with referents of different numbers and genders. Here, too, then, the likelihood of participial agreement can be correlated with the difficulty of processing the sentence.

Hierarchy c). First- and second-person clitics are unambiguously deictic, the first-person form always denoting the speaker and the second-person form the addressee. It is clear that in these cases, the referent can almost always be recovered pragmatically, and the functionality of participial agreement is therefore minimal. Similarly, the referent of a reflexive pronoun is by definition identical with the subject of the verb, and is therefore automatically recoverable from the context. It is with non-reflexive third-person clitics that agreement will be most functional, and it is therefore not surprising that we find it maintained longest in these contexts.

Hierarchy d). In varieties of Romance in which the likelihood of agreement with a third-person clitic-pronoun direct object is dependent on the number and gender of this item (some varieties of Catalan, Sardinian, and Rhæto-Romance), the gender of the singular pronoun preceding the auxiliary 'have' can be retrieved only through the inflection of the participle, whilst the gender in the plural is evident from the surface form of the pronoun itself. In these circumstances, agreement with the object pronoun is functional in the singular, but redundant in the plural; it is therefore not surprising that it should be more prevalent in the former case. It is less easy to account for the apparently differential disappearance of agreement from the plural; but it may be that feminine plural agreement

is more resilient under the influence of agreement with the feminine singular.

Finally, we may note that, in the languages from which object-participle agreement has disappeared completely (Spanish, Portuguese, Rumanian), the third-person accusative clitic pronouns have distinct forms for each number and gender when they occur as the direct object of a compound past tense. In these circumstances, object-participle agreement is redundant, as it conveys no information which cannot be obtained from other items in the sentence.

4.3 *Markedness and the morphosyntactic hierarchies*

The concept of 'markedness' is not unproblematic; and, in many cases (although not all), the choice of one or other member of an opposition as the marked term can appear arbitrary. If we are to make use of the notion, we should certainly try to define it in an independently motivated way; Timberlake's suggestion that "the concept of markedness, or naturalness, must be understood with reference to the particular change involved" (Timberlake 1977: 169), although echoed by subsequent work in natural morphology (compare the notion of 'system-dependent naturalness' developed by Wurzel [1984] and Dressler [1985]), can all too easily lead to circularity. Unmarkedness is defined by a number of well-known criteria, summarized by Battistella (1990: 26) as "optimality, breadth of distribution, syncretization, indeterminateness, simplicity, and prototypicality". Optimality refers to the fact that "When certain segments or certain feature values imply others in language after language, those values are taken to be unmarked" (1990:26). As far as distribution is concerned, "Unmarked terms are distinguished from their marked counterparts by having a greater freedom of occurrence and a greater ability to combine with other linguistic elements" (1990:26) — the characteristic referred to by Croft (1990: 77) as "versatility". The unmarked term is also the one that occurs in positions of absolute neutralization. Syncretization means that "Unmarked categories tend to be more differentiated than marked ones" (1990:27). By the criterion of simplicity "unmarked elements are less elaborate in form than their [marked] counterparts", and by that of prototypicality, they are "experientially more basic" (1990:27). In addition, there is a quantitative yardstick, which is clearly related to some of these qualitative criteria: the unmarked term of an opposition occurs more frequently than its marked counterpart. If all the criteria agree, there will be no problem; if they conflict, it will sometimes be difficult to tell which member of an opposition is marked.

Object-participle agreement in Romance does seem to conform in some respects to the view of markedness put forward by Andersen (1990: 10). For

instance, Nicoli (1983: 371-372), in an aside in his grammar of the Milanese dialect, claims that agreement with a following direct object in standard Italian is more likely to occur in the *trapassato remoto* (or 'past anterior', which, with the auxiliary in the *passato remoto*, or simple past, is both morphologically and stylistically marked) than in other compound past tenses. However, as regards the implicational hierarchies defined in the previous section, I suggest that markedness is a less explicit and less explanatory notion than the sentence-processing hypothesis developed above, and therefore that, if we are to choose between the two accounts, then the latter is to be preferred. Of course, if there is no conflict between a perceptual account and one based on markedness, then the two may be regarded as complementary — we may accept both factors as contributing, at least potentially, to the pattern of change, and are not obliged to choose between them. But it is not clear that the two hypotheses can account equally well for the data. There is little problem with hierarchy a) (the *règle de position*); preceding direct objects are clearly more marked than following direct objects in a language with canonical VO order. Hierarchy d) also accords with markedness principles as far as gender is concerned, since agreement is more frequent in the (marked) feminine, but, when we turn to number, we find that it conflicts with them, in as much as agreement is less likely in the (marked) plural.

Some other aspects of the hierarchies seem even more problematic for a theory of change based on markedness.

- With reference to hierarchy b), it is not obvious that topics, interrogatives, and exclamatives are less marked than relatives. Here, the criteria to some extent conflict. Relatives are more frequent than the other types of preceding direct object; however, unlike topic, interrogative, and exclamative structures, relative clauses are almost by definition subordinate clauses, and, according to this criterion, marked environments (compare the discussion of word-order change in §1).
- Nor is it clear, in respect of hierarchy c), that reflexive third-person clitic-pronouns are less marked than their non-reflexive counterparts. According to qualitative criteria, the reflexive forms are more marked: they exhibit gender and number syncretism; they are defective, in that they lack a subject form; and they are less 'versatile', in that their reference is highly restricted (they must be coreferential with the subject of the verb). However, reflexive forms are more frequent than non-reflexive forms,² and might therefore be regarded as quantitatively unmarked.
- More serious, in respect of the same hierarchy, is the fact that third-person clitic pronouns are arguably less marked than their first- and second-person

equivalents. Greenberg (1966: 84-85) examines data from a number of languages which “lead one to posit, tentatively at least, a hierarchy in which the third person [is] the least marked, and the second person the most marked, with the first person intermediate”. The behavior of the Romance pronouns supports this view — the first- and second-person forms exhibit syncretism of masculine and feminine gender and of accusative and dative case, whilst the third-person forms show greater ‘versatility’ (for instance, they may be anaphoric or deictic. Yet it is precisely with the third-person forms that agreement is most resilient. In these cases, the prediction made by a sentence-processing account and the prediction made by a markedness account are at odds, and it is the former which turns out to be correct. An alternative, discourse-based, view might be put forward, in which discourse participants are less marked than non-participants, and in which the speaker, as the necessary participant in every utterance, is less marked than the hearer, yielding a hierarchy ‘first person > second person > third person’. But, even if this second analysis can be maintained, it will be at best a Pyrrhic explanation: we shall have salvaged a markedness account of the Romance phenomena, whilst demonstrating the Protean nature of markedness.

- Finally, hierarchy b) provides a further problem for a universal definition of markedness. Timberlake (1977: 156) argues that pronouns are less marked than nouns with respect to the changes in Finnish participial complement clauses. Yet agreement in Romance between a participle and a direct object which is a noun disappears earlier than agreement between a participle and a pronoun direct object. The ‘pronoun > noun’ markedness hierarchy is therefore either invalid or not universal.

4.4 *Genre, medium, and style*

Lastly, we should discuss to what extent the distribution of object-participle agreement conforms to the genre, medium, and style hierarchies adumbrated by Andersen (1990:10). In other words, is there a correlation between the markedness of the non-grammatical context and the likelihood that agreement will take place?

This may well be the case for genre in at least some Romance languages. Discussing the decline of object-participle agreement in Old Spanish, Macpherson (1967: 253-254) notes that agreement disappears from prose texts earlier than it does from verse; and, in his survey of the phenomenon in Old Italian, Lucchesi (1963: 255) observes that agreement is more likely to be maintained in

Dante's *Divina Commedia* than in the same author's prose works. As far as Old French is concerned, Busse (1882: 20, 25, 33) finds that agreement is less frequent in popular poetry than in learned verse during the twelfth century; but Wehlitz (1887) maintains that no such distinction exists in the thirteenth.

As for medium, there is evidence that, in French, agreement is lost earlier in speech and retained longer in writing. French authorities as diverse as Foulet (1968: 105) and Tesnière (1965: 581) have claimed that the *règle de position* is essentially bookish, whilst for Price (1971: 233), it is "an artificial rule that is widely ignored in speech and — even if only by inadvertence — occasionally in the written language as well". Of course, this is to some extent a reflection of the inaudibility of agreement with many past participles, as opposed to its systematic visibility in writing (see the discussion in Smith [1993a]). None the less, agreement is audible with a number of past participles; and the impression that we are dealing with a genuine difference between the spoken and the written language is reinforced by the fact that the discrepancy appears to be more evident when the direct object is higher on hierarchy b) — Dauzat (1926: 108), Muller (1964: 44), Bonnaud (1984: 26), and Hagège (1987: 39), for instance, all specifically claim that agreement with a preceding relative (as opposed to agreement with a preceding clitic pronoun) is less likely in spoken French. On the other hand, Alcover (1908: 126) claims that written Catalan exhibits less object-participle agreement than spoken Catalan. Alcover, who has a nationalistic axe to grind, attributes the apparent markedness reversal in this case to Castilian influence, a proposal which is open to question (see Smith 1995a).

However, the evidence regarding loss of agreement in casual, as opposed to formal, style is not so easy to interpret. As early as 1576, the French grammarian, Antoine Cauchie, defined non-agreement as "populaire" and agreement as "plus érudit" (see Rickard 1968: 279). But Cohen (1963: 224ff) claims, albeit anecdotally, that participial agreement in modern French tends to thrive amongst the working class and be slightly less common in the speech of intellectuals; and, according to the survey of Italian undertaken by Hall (1958: 97), teachers are less likely to make the agreement with a preceding relative direct object, whilst white-collar workers are more likely to do so. (However, the reactions of other professions are inconclusive, and Hall's sample of 46 informants is perhaps too small to enable any worthwhile conclusions concerning social distribution to be made.) Brinker (1984: 243-247), on the other hand, in a more recent and thorough survey of Italian, finds that, on the whole, speakers with a higher level of education are more likely to exhibit participial agreement. It is clear that more information is needed, and that extensive sociolinguistic surveys of this phenomenon need to be undertaken before any valid generalizations can be made.

5. Conclusion

In this short paper, I have merely sketched an agenda for future research. Much of the evidence I have cited is impressionistic or anecdotal, and more detailed work needs to be carried out on the issues outlined here before any solid claims can be made. None the less, I hope to have demonstrated two points. First, the claim that linguistic change is realized first in unmarked contexts is anticipated by Bréal, some time before the actual term 'markedness' is coined. Second, the application of theories of markedness to morphosyntactic change is often Procrustean. The view that the actualization of morphosyntactic change is gradual, and proceeds according to a set of hierarchies, and that these hierarchies are determined by factors relevant to the change in question, is undoubtedly well founded; but the claim that they are markedness hierarchies is either uninteresting (if the markedness is defined as purely local) or unsubstantiated (if the markedness is to be defined in terms of universal or quasi-universal principles). An examination of the data concerning object-participle agreement in Romance leads me to conclude that the relation of hierarchies of actualization or extension to a general theory of markedness is perhaps best regarded as epiphenomenal rather than phenomenal.

Author's address

John Charles Smith
St. Catherine's College
Oxford, OX1 3UJ, UNITED KINGDOM
johncharles.smith@st-catherines.oxford.ac.uk

1. Coined, according to Andersen (1989:21) in the 1930s, in discussions between Jakobson and Trubetzkoy. See also Greenberg (1966:62-3), who traces the first use of the term in phonology to a 1931 paper by Trubetzkoy, and its first use in relation to grammatical categories to a 1932 paper by Jakobson, also noting that "earlier adumbrations of these ideas in reference to inflectional categories are to be found in certain Russian grammarians".
2. See the frequency counts, based on 500,000-word corpora, of Juilland & Chang-Rodriguez (1964) for Spanish (8038 *se*, as opposed to 6555 non-reflexive third-person conjunctive pronouns: ratio 1.23:1); Juilland, Brodin & Davidovitch (1970) for French (4637 *se*, as against 4279 non-reflexive third-person conjunctive pronouns: ratio 1.08:1); and Juilland & Traversa (1973) for Italian (5618 *si*, as opposed to 4333 non-reflexive third-person conjunctive pronouns: ratio 1.30:1).

References

- Alcover, A. M. 1908. "Concordansa del participi ab el terme d'acció". *Primer congrés internacional de la llengua catalana, Barcelona, octubre de 1906*, 124-128. [Barcelona]: Horta.
- Andersen, Henning. 1989. "Markedness Theory – The first 150 years". *Markedness in Synchrony and Diachrony* ed. by Olga Mišeska Tomić, 11-46. Berlin & New York: Mouton de Gruyter.
- . 1990. "The Structure of Drift". *Historical Linguistics 1987: Papers from the 8th International Conference on Historical Linguistics. (Lille, 31 August-4 September 1987)* ed. by Henning Andersen & Konrad Koerner, 1-20. Amsterdam: John Benjamins.
- . 1997. "Linguistic Change and the Theory of Markedness". Unpublished paper given at Thirteenth International Conference on Historical Linguistics, Düsseldorf, August 1997.
- Battistella, Edwin L. 1990. *Markedness: The evaluative superstructure of language*. Albany, NY: State University of New York Press.
- Bonnaud, Pierre. 1984. *Mémento abrégé des trois types grammaticaux fondamentaux de l'auvergnat*. Clermont-Ferrand: CNDP-CRDP (Annales du CRDP de Clermont-Ferrand — Documents régionaux).
- Bréal, Michel. 1897. *Essai de sémantique: science des significations*. Paris: Hachette.
- Brinker, Jacques H. 1984. *Problemi dell'accordo del participio passato nell'italiano moderno*. Groningen: Regenboog.
- Busse, Johannes. 1882. *Die Congruenz des Participii Praeteriti in activer Verbal-construction im Altfranzösischen bis zum Anfang des XIII. Jahrhunderts*. Göttingen: Kaestner (Inaugural-Dissertation, Georg-Augusts-Universität, Göttingen).
- Cohen, Marcel. 1963. *Nouveaux regards sur la langue française*. Paris: Éditions sociales.
- Croft, William. 1990. *Typology and Universals*. Cambridge: Cambridge University Press.
- Dauzat, Albert. 1926. *La langue française: sa vie, son évolution*. Paris: Stock.
- Dressler, Wolfgang U. 1985. *Morphonology: The dynamics of derivation*. Ann Arbor: Karoma.
- Foulet, Lucien. 1968. *Petite syntaxe de l'ancien français*. Paris: Champion.
- Greenberg, Joseph H. 1963. "Some Universals of Grammar With Particular Reference to the Order of Meaningful Elements". *Universals of Language* ed. by Joseph H. Greenberg, 58-90. Cambridge, Mass.: M.I.T. Press.
- . 1966. "Language Universals". *Current Trends in Linguistics: Vol. III, Theoretical Foundations* ed. by Thomas A. Sebeok, 61-112. The Hague & Paris: Mouton.
- Hagège, Claude. 1987. *Le français et les siècles*. Paris: Odile Jacob.
- Hall, Robert A., Jr. 1958. "Statistica sintattica: l'accordo del participio passato coniugato con avere". *Lingua nostra* 19.95-100.
- Harris, Alice C. & Lyle Campbell. 1995. *Historical Syntax in Cross-Linguistic Perspective*. Cambridge: Cambridge University Press.

- Juilland, Alphonse & Eugenio Chang-Rodriguez. 1964. *Frequency Dictionary of Spanish Words*. The Hague: Mouton.
- Juilland, Alphonse, Dorothy Brodin & Catherine Davidovitch. 1970. *Frequency Dictionary of French Words*. The Hague & Paris: Mouton.
- Juilland, Alphonse & Vincenzo Traversa. 1973. *Frequency Dictionary of Italian Words*. The Hague & Paris: Mouton.
- Lucchesi, Valerio. 1963. "L'accordo fra participio passato e oggetto nei tempi perifrastici retti da "avere" nel volgare antico (secc. XIII-XIV)". *Atti e memorie dell'Accademia toscana di scienze e lettere* 27.191-278.
- Macpherson, Ian R. 1967. "Past Participle Agreement in Old Spanish: Transitive verbs". *Bulletin of Hispanic Studies* 44.241-254.
- Matsuda, Kenjiro. 1997. "On the Conservatism of Embedded Clauses". Unpublished paper given at Thirteenth International Conference on Historical Linguistics, Düsseldorf, August 1997.
- Montes Giraldo, J. J. 1976. "Un arcaísmo gramatical en Colombia: la construcción del pretérito compuesto de subjuntivo con *ser*". *Thesaurus (Boletín del Instituto Caro y Cuervo)* 31.51-52.
- Muller, Charles. 1964. "Les français et leurs participes". *Le français dans le monde* [avril-mai 1964], 43-44.
- Nicoli, Franco. 1983. *Grammatica milanese*. Busto Arsizio: Bramante.
- Price, Glanville. 1971. *The French Language: Present and past*. London: Edward Arnold.
- Rickard, Peter. 1968. *La langue française au seizième siècle: Étude suivie de textes*. Cambridge: Cambridge University Press.
- Ronjat, Jules. 1937. *Grammaire istorique [sic] des parlers provençaux modernes: Tome III*. Montpellier: Société des langues romanes.
- Roux, J. 1895. *Grammaire limousine*. Brive: École limousine.
- Seifert, Eva. 1935. *Tenere 'haben' im Romanischen*. Firenze: Olschki.
- Sicre, Paul. 1909. *Éléments de grammaire du dialecte de Foix*. Foix: Gadrat aîné.
- Smith, John Charles. 1989. "Actualization Reanalyzed: Evidence from the Romance compound past tenses". *Synchronic and Diachronic Approaches to Linguistic Variation and Change (Georgetown University Round Table on Languages and Linguistics, 1988)* ed. by Thomas J. Walsh, 310-325. Washington, D.C.: Georgetown University Press.
- . 1991. "Problemi dell'accordo del participio passato coll'oggetto diretto nei tempi composti coniugati con avere in italiano, con speciale riferimento ai dialetti". *Tra Rinascimento e strutture attuali: Saggi di linguistica italiana (Atti del Primo Convegno della Società Internazionale di Linguistica e Filologia Italiana, Siena, 28-31 marzo 1989, volume I°)* ed. by Luciano Giannelli, Nicoletta Maraschio, Teresa Poggi Salani & Massimo Vedovelli, 365-371. Torino: Rosenberg & Sellier.
- . 1993a. "The Agreement of the Past Participle Conjugated with avoir and a Preceding Direct Object: A brief history of prescriptive attitudes". *Authority and the*

- French Language: Papers from a conference at the University of Bristol* ed. by Rodney Sampson, 87-125. Münster: Nodus.
- . 1993b. "La desaparición de la concordancia entre participio de pasado y objeto directo en castellano y catalán: aspectos geográficos e históricos". *Actas del Primer Congreso Anglo-Hispano: tomo I, Lingüística* ed. by Ralph Penny, 275-285. Madrid: Castalia.
- . 1995a. "Agreement between Past Participle and Direct Object in Catalan: The hypothesis of Castilian influence revisited". *Linguistic Change under Contact Conditions* ed. by Jacek Fisiak, 271-289. Berlin & New York: Mouton de Gruyter.
- . 1995b. "Perceptual Factors and the Disappearance of Agreement between Past Participle and Direct Object in Romance". *Linguistic Theory and the Romance Languages* ed. by John Charles Smith & Martin Maiden, 161-180. Amsterdam & Philadelphia: John Benjamins.
- . 1996. "Surfonctionnalité et hyperanalyse: l'accord du participe passé dans les langues romanes à la lumière de deux théories récentes". *Faits de langues* 8.113-120.
- . 1997. "Types and Tokens in Language Change: Some evidence from Romance". *Language History and Linguistic Modelling: a Festschrift for Jacek Fisiak on his 60th birthday* ed. by Raymond Hickey & Stanislaw Puppel, 1099-1111. Berlin & New York: Mouton de Gruyter.
- Tesnière, Lucien. 1965. *Éléments de syntaxe structurale*. 2nd ed. Paris: Klincksieck.
- Timberlake, Alan. 1977. "Reanalysis and Actualization in Syntactic Change". *Mechanisms of Syntactic Change* ed. by Charles N. Li, 141-177. Austin & London: University of Texas Press.
- Wehlitz, Hermann. 1887. *Die Congruenz des Participii Praeteriti in activer Verbalconstruction im Französischen (vom Anfang des XIII. bis zum Ende des XV. Jahrhunderts)*. Greifswald: Abel. (Inaugural-Dissertation, Universität Greifswald.)
- Wheeler, Max W. 1988. "Catalan". *The Romance Languages* ed. by Martin Harris & Nigel Vincent, 170-208. London: Routledge.
- Wurzel, Wolfgang U. 1984. *Flexionsmorphologie und Natürlichkeit*. Berlin: Akademie-Verlag.

CHAPTER 36

Romanian and the Balkans

Some comparative perspectives

Brian D. Joseph

The Ohio State University

1. Introduction

Comparison is at the heart of linguistic methodology, especially as concerns historical linguistics. It was the development of the comparative method, for instance, which made possible the great advances in the 19th C. that E. F. Konrad Koerner has done such a service to the field by documenting and analyzing in his many works on the history of linguistics (note especially the articles on August Schleicher, Franz Bopp, Friedrich Schlegel, and Jacob Grimm — Koerner 1982, 1984, 1987, and 1988, respectively, all reprinted in revised and augmented form in Koerner 1989). In the spirit of recognizing his own many contributions, I offer this contribution to comparative linguistics, focusing on Romanian.

Romanian occupies a position that makes it of particular interest to linguists with comparative interests, for it allows for several crucial types of comparisons to be made. First, comparisons of a genetic nature are possible, pertaining to Romanian in the context of those languages most closely related to it, and thus looking at Romanian vis-à-vis the other Romance languages. Second, comparisons can be made that are of a geographical nature, pertaining to Romanian in the context of its linguistic neighbors, and thus looking at it in relation to the other languages of the Balkans. These two types of comparison are in addition to general typological comparisons that can be made between any two (or more) languages, in the interests of pursuing linguistic universals and determining the nature of Universal Grammar.

In a sense, just as general typological comparisons can take any language(s) as a starting point, so too can the first two types of comparison, genetic and geographic, be made for any language. In particular, with the exception of the relatively few generally agreed-upon language isolates, e.g., Basque or Sumerian,¹ there will always be genetic relatives of a language with which to make comparisons.² Similarly, there are always linguistic neighbors somewhere within or alongside the territory in which a given language is spoken, even if only at its outer reaches.

What makes Romanian of special interest in this regard is the fact that its closest relatives — the members of the Romance group of languages — are so well-known and so well-studied that illuminating comparisons can be made on matters of linguistic detail that might not be possible within other genetic groupings that are less well-documented. As for its geographic setting, Romanian again is of special interest since along with most of its neighbors in the Balkans, it participates in numerous areally based similarities that define the Balkan *Sprachbund*,³ an area where long-term intense and intimate contact among speakers of several different languages has led to massive structural convergence.

The languages that participate to some significant extent in the Balkan *Sprachbund* convergences can be referred to as “Balkan languages”⁴ and besides Romanian include Albanian (both major dialects: Geg (North) and Tosk (South) but especially Tosk); Bulgarian; Greek; (Slavic) Macedonian;⁵ Romany (the language of the Indic Gypsies), Serbo-Croatian⁶ (with the Torlak dialects of Southeast Serbia being most relevant), and Turkish (though mostly relevant for its lexical contributions). Actually, in counting Romanian among the Balkan languages, the greatest attention belongs to Aromanian, spoken in pockets in northern and central Greece, in Albania, and in Macedonia, and to Megleno-Romanian, spoken in a few areas in northern Greece; Daco-Romanian is a Balkan language to some extent, but not fully so, and Istro-Romanian is largely irrelevant as far as the Balkan *Sprachbund* is concerned. The extent of participation in the *Sprachbund* for some of these Romanian varieties⁷ is discussed below in section 3.

There are many features on which various of the Balkan languages agree and have converged,⁸ covering all levels of linguistic structure, but among the most salient and most widely discussed ones are those listed in (1) to (3); for a fuller discussion, the reader is referred to various handbooks such as Sandfeld 1930, Schaller 1975, and Banfi 1985, among others:⁹

- (1) a. Balkan Sprachbund phonological features
 - absence of “overlay” features, e.g., length and nasalization, in the articulation of vowels
 - presence of a mid-to-high central vowel
- b. Balkan Sprachbund morphological features
 - enclitic definite article
 - invariant future tense marker derived from a verb meaning “want”
 - merger of genitive and dative case
 - analytic comparative
- c. Balkan Sprachbund syntactic features
 - use of a special verb form to indicate confirmativity/evidentiality
 - cross-indexing (“doubling”) of direct and indirect objects by a weak (“clitic”) pronoun
 - use of finite complementation in place of infinitives

These features are not realized in all of the Balkan languages nor are they necessarily found uniformly in the languages that show them to any extent; still, even if no single one is criterial for membership in the *Sprachbund*, the overlapping clusters of these features in the various languages serve to define the *Sprachbund*.

Recognizing the Balkan *Sprachbund*, then, allows for comparisons of a particularly interesting sort to be made for Romanian, taking in not just genetic and typological comparisons but also those having to do with the special character of language contact in the Balkans and thus with what can be revealed about the historical development of Romanian on its own and in relation to the other Balkan languages.

Accordingly, several instances of comparisons involving Romanian and its Balkan neighbors, especially Greek and Albanian, are discussed here. Inasmuch as the most interesting convergences in the Balkans involve morphology and syntax (see n.9 regarding some of the putative phonological Balkanisms), the examples presented here focus on (morpho-)syntax. These comparisons are undertaken with an eye toward illuminating which are of interest for the purely historical purpose of determining the diachrony of Romanian and its neighboring languages, which are of relevance to understanding the nature of language contact in the Balkans and in general, and which have significance for furthering our knowledge of Universal Grammar. These various goals for comparison cut across the three types of comparison discussed above: genetic, geographic, and typological, and thus can draw on all three.

2. A possible substratal effect in Romanian syntax

The first example to be discussed involves a syntactic anomaly within Romanian prepositional syntax that is paralleled in Albanian.¹⁰ Somewhat idiosyncratically for prepositions in Romanian, which normally allow only an indefinite noun phrase as object, as in (2):

- (2) a. *din cladire*
 from building/INDEF
 “from a/the building”
 b. **din cladirea*
 building/DEF
 “from the building”

the preposition *cu* “with”, requires a definite article for an unmodified nominal object and allows a definite article for a modified object:

- (3) a. *cu prietenul*
 with friend/DEF
 “with the friend”
 b. **cu prieten*
 friend/INDEF
 “with a friend”
 c. *cu un prieten bun*
 a friend/INDEF good
 “with a good friend”
 d. *cu prietenul bun*
 friend/DEF good
 “with the good friend”

While there is no apparent synchronic reason for this special behavior of *cu*, so that most likely it requires lexical stipulation as being exceptional to the general pattern of only indefinite objects with prepositions, the comparison with Albanian, especially when taken together with some other facts about Romanian and Albanian, suggests a historical reason for the anomaly, which moreover sheds some light on the historical antecedents of Romanian itself.

In particular, in Albanian also, the preposition for “with”, *me*, is unusual in requiring a definite object, even though the object of most prepositions can occur with the definite article or not, according to meaning. Thus the contrast between *pa* “without” in (4) and *me* “with” in (5) is instructive in this regard:

- (4) a. *pa sheqer*
without sugar/INDEF
“without sugar”
b. *pa mjetet*
without tools/DEF
“without the tools”
- (5) a. *me shqiptarët*
with Albanians/DEF
“with (the) Albanians”
b. **me shqiptarë*
Albanians/INDEF

Moreover, as in Romanian, Albanian allows indefinite objects with *me* when the noun is modified:

- (6) *një lëvizje me karakter politik*
a movement with character/INDEF political
“a movement with a political character”

There may well be some differences in the range of nouns for which use of the definite article after *me* and *cu* is required or impossible, as suggested by Boretzky 1968. Still, this agreement that Albanian and Romanian show with regard to a synchronic anomaly involving prepositions and definiteness is striking and therefore quite intriguing, and is something that demands an explanation at least in historical terms even if an illuminating synchronic account is not possible.

This parallel seems to be too precise to be due to chance, and it certainly cannot be a matter of lexical borrowing since the form of the preposition is different in each language (*cu* vs. *me*). It is possible that it is a structural borrowing from one language into the other or even a calque (loan-translation), but most of the structural parallels that are found between Romanian and Albanian are ones that are shared with other Balkan languages (e.g., regarding the infinitive or the falling together of dative and genitive cases) and thus are more likely to be the result of the various processes that led to overall convergence in the Balkan languages, or else have a Latin origin (thus inherited in Romanian and the result of heavy Latin influence on early Albanian). While it has been suggested (e.g., by Reichenkron 1962) that there is a common Romance source for the prepositional parallel under consideration here, one other possibility presents itself as a compelling source for this Romanian-Albanian convergence, namely that of spread from a substratum.

The spread of substratum elements into a second language presumably takes place through a period of language shift, as speakers of one language carry over traits of their primary language into the language they are shifting to. Although substratum effects are real enough, as is evident from observations of second-language acquisition, whether in tutored (e.g., academic) environments or in a natural untutored learning situation, substratum accounts in historical linguistics are often hard to maintain since they generally seem to be little more than a last resort "explanation" rather than anything substantive.¹¹ Still, Thomason & Kaufman 1988 have argued that in real cases of substratum interference, one typically finds effects in both syntax and phonology, and this observation then provides a basis for judging this particular Romanian-Albanian convergence, since overall the special parallels between these two languages fit this pattern fairly well.

In particular, in addition to this syntactic parallel involving certain prepositional objects, there is other evidence for the view expressed by Hamp (1989: 47) that "historically Romanian is Latin spoken with an Albanian stress system", i.e., with the "Danubian Late Latin" of Dacia being filtered through the grammars of one group of Proto-Albanian (what Hamp sometimes calls "Albanoid") speakers. This evidence consists of striking parallels in stress placement and stress shift, whereby both languages have stress on the final syllable of the stem of the lexeme, so that the stress shifts to the right when a derivational suffix is added, e.g., agentive *-tor-* in Albanian or deverbal adjectival *-tor-* in Romanian, but not when an inflectional ending, e.g., Albanian dative plural *-eve* or Romanian definite dative singular *-lui* or feminine singular *-e* is added:

- (7) Albanian
katúnd "village/NOM.SG" / *katúndeve* "to villages/DAT.PL"
púnë "work/NOM.SG" / *punëtór* "worker/NOM.SG" / *punëtóreve*
 (DAT.PL)
- (8) Romanian
cîtne "dog/NOM.SG" / *cînelui* "to the dog/DEF.DAT.SG"
a folosi "to use" / *folositor* "useful/MASC.SG" / *folositoare* (FEM.SG)

Thus these striking parallels can be explained as the result of substratum influence in Romanian, as "Albanoid"/Proto-Albanian speakers shifted to the Latin spoken in Dacia and carried these aspects of their native syntax and prosodic system into emerging Romanian. The fact then that there is both an accentual parallel and a syntactic parallel is highly suggestive of a substratum, and the fact that only two languages are involved — Romanian as the shifted-to

language and (modern) Albanian as the sibling language of that spoken by the shifting population — makes a substratum scenario quite plausible. Thus while a synchronic explanation for the definiteness effect with *cu* and *me* may still be lacking, some historical perspective is possible, thus enriching our understanding of the emergence of Romanian and its prehistoric connections with Albanian (or “Albanoid”) speakers. The comparisons made here with Albanian, both as to the primary focal point regarding “with” phrases and to the secondary comparison made regarding accent rules, therefore together serve to illuminate an otherwise obscure aspect of Romanian prepositional phrase syntax.

3. Comparisons within a Sprachbund and a Romance context

The second goal given above for comparison, namely extending our understanding of the Balkan Sprachbund and thus shedding light on the nature of language contact in the Balkans, emerges from the parallel seen in the formation of compound tenses in Romanian and other Balkan languages, especially in regard to perfects formed with “have” as an auxiliary. A consideration of this feature provides some evidence bearing on the question of just how “Balkan” Daco-Romanian is, as opposed to other varieties of Romanian, and in particular Aromanian and Megleno-Romanian, and also clarifies the place of different varieties of Romanian in an overall Romance context. In this way, therefore, some perspective is gained on the characterization of Daco-Romanian as a Romance language as opposed to a Balkan language.¹²

At issue is the composition and value of a perfect tense formation made up of “have” and a participle that independently functions as a passive participle to transitive verbs and an active participle to intransitives. Of particular interest is the fact that in Aromanian and Megleno-Romanian, as in Albanian, Greek, and Macedonian, there is a past perfect formation, utilizing the past of “have” as the auxiliary verb preceding the participle, as exemplified in (9):

- | | | |
|-----|-------------------|--------------------------------------|
| (9) | Albanian: | <i>kisha lidhurë</i> “I had tied” |
| | Greek: | <i>íxa ðeméno</i> “I had tied” |
| | Macedonian: | <i>imav storeno</i> “I had made” |
| | Aromanian: | <i>avea mîcatà</i> “(s)he had eaten” |
| | Megleno-Romanian: | <i>vea durmit</i> “(s)he had slept” |

While other forms of such a “have” perfect system can be found, e.g., a present perfect such as Greek *éxo ðeméno* “I have tied”,¹³ it is the past perfect in general that seems to be historically prior — and is demonstrably so in Greek at least

(see Joseph 1983, 1998) — and which has been the focus of Balkanological interest.¹⁴

In Daco-Romanian, by contrast, the only relevant comparable verbal formation, consisting of an auxiliary form of “have”¹⁵ and a participle with a similar valency to the Greek, Albanian, etc. participle, is the *perfectul compus* (“compound perfect”), a full paradigm for which is given in (10) for the verb *a lucra* “to work”:

(10)	1SG <i>am lucrat</i>	1PL <i>am lucrat</i>
	2 <i>ai lucrat</i>	2 <i>ați lucrat</i>
	3 <i>a lucrat</i>	3 <i>au lucrat</i>

This tense has a similar form to the Balkan perfect in (9), though admittedly with an apparent present form of the auxiliary,¹⁶ but importantly, in general it instead has the meaning of a simple past, and not a perfect; thus, the forms in (10) mean “I worked; you worked, (s)he worked, etc”. and not “I have worked, etc”. In this way, the Daco-Romanian “have” periphrastic tense is more like the French *passé composé*, e.g., *j’ai travaillé* “I worked”, which has ousted the simple (monolectal) preterite, the so-called *passé simple*, from “normal spoken French” (Harris 1987: 221),¹⁷ and it also shows an affinity, moreover, with relatively recent developments in Castilian Spanish (so Green 1987: 257) and in northern Italian dialects (Vincent 1987: 279), in which a compound past tense that originated as a “have”-perfect has encroached upon the function of a simple preterite. In that way, therefore, Daco-Romanian seems more like a proto-typical member of the Romance language group than a member of the Balkan Sprachbund, in that it is taking part in this widespread Romance drift concerning the value of various tense formations.¹⁸

Thus Daco-Romanian really does not match up well with the other Balkan languages with regard to showing evidence of a “have” perfect system, since it has a formation that is superficially somewhat parallel in terms of its form but quite different in terms of its meaning. Moreover, when approached from the perspective of meaning, it is noteworthy that the Daco-Romanian form corresponding in meaning to a past perfect is not a periphrastic tense at all, but rather is a synthetic, i.e., monolectal, formation, as in (11):

(11)	1SG <i>lucasem</i>	1PL <i>lucaserăm</i>
	2 <i>lucaseși</i>	2 <i>lucaserăți</i>
	3 <i>lucase</i>	3 <i>lucaseră</i>
	“I had worked, you had worked, etc.”	

Again, this situation seems more characteristic of Romance languages in general;

Portuguese, for instance, has a synthetic past perfect (e.g., *falara* "I had spoken") and a periphrastic perfect (with the verb *ter* "have" as the auxiliary, e.g., *tenho tomado* "I have taken").¹⁹ Furthermore, in the one part of what might be viewed as a perfect "system" in Daco-Romanian in which one finds a periphrastic formation, namely the future perfect, the auxiliary is the verb "be", *a fi*, and not "have", as in (12):²⁰

- | | | | |
|------|--------------------------------------------------------------------------|---|----------------------------------------------------------------------------------|
| (12) | <i>voi fi lucrat</i>
will/1SG be/INF work/PPL
"I will have worked" | / | <i>vom fi lucrat</i> (etc.)
will/1PL be/INF work/PPL
"we will have worked" |
|------|--------------------------------------------------------------------------|---|----------------------------------------------------------------------------------|

Such a use of "be" in active forms within a system that can be characterized as marking "perfect" (as opposed to perfective aspect) is reminiscent of the Romance (e.g., French and Italian) distinction between "have" and "be" active compound tenses (e.g., French *j'ai travaillé* "I worked", with "have", versus *je suis tombé* "I fell" with "be") but more likely represents the result of influence from surrounding Slavic languages, especially Bulgarian; compare (13) with (12):

- (13) *šte süm pisał*
 FUT am/1SG write/PPL
 "I will have written".

Still, of the varieties of Romanian other than Istro-Romanian, Daco-Romanian is the least like the core Balkan languages — and the most like other Romance languages — while Aromanian and Megleno-Romanian are far more Balkan in their character, at least insofar as the perfect and the preterite tenses are concerned. Affinities that Daco-Romanian may show with Balkan languages in this regard (cf. (12) and (13)) have not been such as to give all parts of its perfect and preterite system an overall Balkan character.

Sandfeld (1930: 105) points out another related way, pertaining to the perfect and the preterite, in which Aromanian, as opposed to Daco-Romanian, shows a more Balkan-oriented, and less Romance-oriented, character. He notes that southern varieties of Aromanian show a preference for the simple preterite over the composite preterite, saying "c'est surtout vrai pour les cas où le préterit simple a le sens d'un 'perfectum praesens'" and assessing this preference as "sans doute sous l'influence de l'aoriste grec".²¹ He cites as an example Aromanian *nu s mârță sor mea* "my sister is not married" (i.e., "has not married", literally "not self married/PRET sister my"), which he contrasts with Daco-Romanian *nu s'a măritat sora mea*, with the *perfectul compus* (and thus literally "not self has/AUX married sister my"). In this way, Aromanian aligns itself with Greek, since in Greek the simple past is far more common than the present

perfect,²² and has been for centuries; indeed, 17th C. Greek grammarians (Germano 1622, Portius 1638) give past perfect forms such as *eíkha grápsei* 'I had written' as the equivalent of a Latin *plus-quam-perfectum* (i.e., pluperfect or past perfect), but list the simple past (corresponding to the Ancient Greek aorist) *égrapsa* 'I wrote; I have written' as the equivalent of a present perfect, and Modern Greek usage confirms this distribution (see Thumb 1912: §229).

Overall, therefore, this exercise in areal and genetic comparisons involving equivalents of perfects and simple preterites reveals that Daco-Romanian is less fully a Balkan language, where "Balkan" marks a language as having taken part in the contact situations that led to the massive structural convergences characteristic of the *Sprachbund*, than Aromanian or Megleno-Romanian, and further that it shows more affinities with other Romance languages in this domain than do either Aromanian or Megleno-Romanian. The comparisons made here thus clarify the place of Daco-Romanian and the other varieties of Romanian both within the Balkans and within the Romance group, and further point to Aromanian and Megleno-Romanian as having taken part more fully in the intense and intimate contact that gave rise to the Balkan *Sprachbund*.²³

4. Universally oriented comparisons

As noted in the introduction, comparisons can also be made on matters of linguistic typology for the purpose of illuminating some aspect of Universal Grammar. Such comparisons are of course possible with any set of two or more languages, whether or not they are genetically or geographically related, and in fact unrelated languages are often preferred, by way of diversifying the sample. Still, Universal Grammar must be broad enough to take in facts from any natural language, and to be sure, many claims about the universal behavior of clitic pronouns, for instance, have been based, at first at least, on their behavior in Romance languages. Thus, comparisons made with languages within the Balkans can also further this goal and indeed, researchers have not shied away from typologically and universally oriented comparisons, including those involving Romanian; one such case is presented here in some detail.²⁴

For example, Joseph (1980a, 1983: 232ff.) noted a parallel between Greek and Daco-Romanian in the realization of *tough*-Movement type sentences, pointing out that in such sentences in each of the languages, the finite, tensed complement clause — the only possibility in Greek and one of two possible forms of the complement in Romanian, the other being the nonfinite supine, shown in (14c) — cannot be both transitive and objectless. In Greek, shown in

(14a), this state is achieved by the occurrence of a resumptive pronominal object in the complement clause corresponding to the moved nominal, and in Romanian, shown in (14b), by the occurrence of a passive complement verb:²⁵

- (14) a. *ta angliká_i íne đískola na*
 the English/NTR.NOM.PL are/3PL difficult/NTR.NOM.PL SUBJUNC
ta_i katalávi kanís
 them/NTR.ACC.PL understand/3SG someone/NOM.SG
 "English is difficult to understand" (literally = "the-English
 (things) are difficult that someone understand them")
- b. *asta nu-i greu să se facă*
 this not-is/3SG hard SUBJUNC REFL do/3SG
 "This is not hard to do" (literally = "This is not hard that it be
 done")
- c. *asta nu-i greu de făcut*
 this not-is/3SG hard do/SUPINE
 "This is not hard to do" (literally = "This is not hard that it be
 done")

This parallel was taken to be a reflection of a universal tendency, evident from cross-linguistic comparisons (Joseph 1978/1990, 1980b), for the complement clause in *tough*-Movement constructions to be nonfinite when objectless, as with the Romanian supine type in (14c) or the English translations in (14). The Romanian evidence, especially when taken together with the Greek,²⁶ was viewed as supporting a putative linguistic universal that *tough*-Movement cannot deprive a finite verb of its object, and provided an important dimension to the investigation of this hypothesized universal constraint by furnishing examples of the construction with a finite complement clause, thus showing that the universal cannot simply be stated as a requirement for a nonfinite complement in this construction. While other research casts doubt on the validity of such a constraint,²⁷ or on the need to single out *tough*-Movement constructions in the statement of the constraint,²⁸ the value of the cross-linguistic comparisons is undeniable, as is the contribution of the Romanian evidence to the investigation of this constraint. Moreover, even if the constraint is not valid, the Balkan-internal parallel that Romanian and Greek (and Albanian — see n.26) offer evidence of is noteworthy.

However, it must be recognized that if the goal of these comparisons is to elucidate some aspect of Universal Grammar, then the actual terms of the comparison, that is the particular languages being compared, are perhaps less important to the investigation than the mere fact of comparison. In a sense, the

role in any such comparisons of Romanian itself and of Romanian in relation to other Balkan languages is really accidental, since the purpose of developing a linguistic typology can be served by comparisons between Romanian and any other, non-Balkan, language, or for that matter, as noted above, by comparisons between any arbitrarily selected languages.

Moreover, if there is a Balkanological goal to the comparison involving Romanian, then it is not clear that anything significant can emerge by comparisons inspired by typological and universalist considerations; what led to a special "Balkan" character to Greek, Albanian, Aromanian, and to a lesser extent Daco-Romanian, is the special language contact situation in the Balkans and there is no reason to believe that Universal Grammar plays a role in language contact except in the most trivial sense that natural languages are involved.²⁹ Thus the account given above of *tough*-Movement structures in the Balkans does not really relate in any way to the Balkan Sprachbund, however (potentially) interesting it might be for understanding the nature of *tough*-Movement and related constructions universally.³⁰

5. Comparison for the purpose of finding differences

Most of the comparisons discussed here have focused on similarities between Romanian and other languages, though in some instances, the comparisons revealed points of difference as well. Since Universal Grammar is as much defined by accounting for the differences among languages as it is by accounting for their similarities, by way of conclusion, an example is provided here of a comparison that leads to the identification of a point of grammar on which Romanian seems to differ from all other languages within the Balkans and within Romance, and quite possibly as well from all other languages period.

The particular feature in question concerns the placement of weak, so-called "clitic", object pronouns. For the most part, weak object pronouns are positioned in Romanian in reference to the inflected verb, generally occurring as proclitics, even in compound tenses, but enclitic to imperatives and gerunds:

- (15) a. *îmi dau cadoul*
 me/DAT give/3PL present/DEF
 "They give the present to me"
- b. *mi-l dau*³¹
 me/DAT it/ACC give/3PL
 "They give it to me"

- c. *mi- l au dat*
me/DAT it/ACC have/3PL given/PPL
"They gave it to me"
- d. *spune- i*
tell/IMPV.SG him/DAT
"Tell him!"
- e. *cerîndu- le*
ask/GER them/ACC
"asking them ..."

In this way, Romanian seems quite parallel to other patterns of clitic placement found in Romance languages, e.g., French, and in the Balkans, e.g., Greek.³²

One idiosyncratic exception to this placement, however, occurs with the feminine singular accusative weak pronoun *o*, and only with that pronoun. What is found in formations such as the *perfectul compus* is that instead of *o* being positioned before the tensed verb, as with *mi* or *l* in (15c), it occurs after the participial part of the compound tense; this contrasts with the placement of *o* before the verb in a simple tense, as in (16c):

- (16) a. *am văzut- o*
have/1SG seen/PPL her/ACC.WK
"I saw her"
- b. **o-am văzut*
- c. *o găsim*
her/ACC.WK find/1PL
"We find her".

Although there may be some phonological issues that play a role in this placement (see Wanner 1993), it is still the case that only the feminine singular accusative weak pronoun shows this behavior. Such an idiosyncratic exception, and particularly one centered on the feminine singular pronoun, is unique within the Balkans and seems also to be unique within the Romance group of languages.³³ Moreover, the highly specific nature of this idiosyncrasy, coming just within the placement of weak pronouns, just for the feminine singular form, and just in formations such as the compound past, is so specialized that it would be hard to imagine that there could be another language outside of Romance and outside of the Balkans that has precisely the same irregularity.

What this example shows, therefore, is that even with the quest for generalizations, especially those that cut across languages and are thus of interest from a universalist perspective or the perspective of language contact, there are still

going to be highly particularized aspects of languages that are real, even if messy. They therefore cannot be ignored. Romanian, through the comparisons explored here, reveals itself as a member to some extent of the Balkan Sprachbund, as a member of the Romance group, and as a member of the set of human languages more generally, but at the same time these comparisons show that it has some unique characteristics that are not paralleled elsewhere.

Author's address

Brian D. Joseph
Department of Linguistics
222 Oxley Hall
The Ohio State University
Columbus, OH 43210-1298, USA
bjoseph@ling.ohio-state.edu

Notes

1. Recent attempts to connect these languages to other known language families notwithstanding, I maintain here the *communis opinio* that these are genetically isolated languages, given our current state of knowledge. See Trask 1997 for some discussion of Basque as an isolate.
2. And even for isolates there are comparisons to be made among different dialects of the language.
3. I use *Sprachbund* as a technical term rather than any of its clumsy and infelicitous possible translations such as "linguistic union" or "language league"; "linguistic area" is sometimes used in English, but "convergence area" probably comes closest to being a suitably apt term.
4. This is intended as a technical term, following the distinction drawn by Schaller 1975 between "Balkan languages", those that are Sprachbund members, and "languages of the Balkans", a purely geographic classification that includes all the Balkan languages but also (not counting very recent migrations, e.g., by Arab speakers into Greece, or adventitious occurrences, such as American English speakers residing in Greece): Armenian (spoken in Bulgaria); Circassian (Adygey variety; spoken in the Kosovo area of [former] Yugoslavia); German (spoken in Romania); Hungarian (spoken in Romania); Italian (spoken in the Istria area of [former] Yugoslavia); Judezmo (the dialect of Spanish spoken in parts of Greece by former Iberian Jews, also known as Judeo-Español or occasionally, and perhaps erroneously, Ladino); Ruthenian (also known as Rusyn, spoken in Vojvodina area of [former] Yugoslavia and perhaps actually a dialect of Ukrainian); Slovak (in a small enclave in Vojvodina area of [former] Yugoslavia); and Slovenian.
5. I say 'Slavic Macedonian' for purposes of clarity (not ideology), in order to distinguish it clearly from Ancient Macedonian, an entirely different language possibly closely related to Greek and spoken in the first millennium B.C. also in the Balkans; sometimes the autonym 'Makedonski' is used in English for the modern Slavic language, with 'Macedonian' being reserved for the ancient language, but no usage has yet gained full acceptance among scholars.

6. Now more usually referred to separately as Serbian and Croatian, as well as Bosnian; I adopt here the still traditional joint designation though the actual labels are not significant for the purposes of this paper.
7. I take no position here on whether or not these varieties represent four distinct but closely related languages or instead are dialects of a single language. While I am inclined toward the former view, nothing crucial in this discussion hinges on that decision, and my use of the term 'Romanian' herein should be taken to mean Daco-Romanian, in the unmarked case.
8. Indeed, one of the most fascinating aspects of the Balkan Sprachbund is that virtually all of the significant structural convergences represent also significant *divergence* from earlier states of each of the languages.
9. There is considerable debate on the validity of several of these features as true *Sprachbund* features; for example, there may not be an adequate language contact scenario to account for the appearance of a mid-to-high central vowel in many of these languages and it does have different historical sources in the languages that show it. Also, some are found outside the Balkans (e.g., Scandinavian languages and Northern Russian dialects have developed a postpositive definite article, and English has a future based on a volitional verb), so that presumably they can develop independently in a given language. Nonetheless, these features are listed here as representative of what is typically cited in the literature.
10. Noted, for example, in Schaller (1975: 169), with some discussion of relevant literature, and in Banfi (1985: §3.6.6); the Albanian facts are taken in part from Newmark, Hubbard & Prifti, 1982.
11. Indeed, some substratum explanations have been given for various Balkan features, but without much credibility. For example, it has been claimed that the loss of the infinitive is due to the workings of a prehistoric Thracian or Illyrian substratum, but the chronology of the replacement of earlier infinitives by finite complementation argues against this; for that was an innovation that took place within the fairly recent (for the Balkans) historical period of the Post-Classical period and on into the Medieval period, inasmuch as infinitives persist into the Medieval period for all of the Balkan languages— see Joseph 1983, especially chapter 7, for discussion.
12. I speak here somewhat figuratively, for the terms of comparison are not equivalent nor are they mutually exclusive: "Romance" is a genetic classification while "Balkan" is rather a geographic and sociolinguistic classification (sociolinguistic in the sense of defining a zone of a special type of contact among speakers). A presupposition here, one which admittedly may not be warranted, is that there is a general "type" that one finds among Romance languages and which therefore is a norm against which all genetically Romance languages can be measured.
13. Greek also shows a perfect formation, an innovation dating to the Medieval Greek period, with "have" followed by an invariant verb form which historically continues the older infinitive, but which synchronically may be nothing more than a variant participial form, e.g., *éxa óési* "I had tied", as discussed in Joseph (1983, 1999b).
14. The Balkan character of this parallel was discussed by Sandfeld (1930: 105–6).
15. The paradigm of "have" in this formation overlaps with that of the present tense of the main verb "have", differing in the 3SG (*a* vs. *are* "(s)he has"), 1PL (*am* vs. *avem* "we have"), and 2PL (*ați* vs. *aveți* "you (all) have"). The forms are historically related, to be sure, so that treating the auxiliary in this formation as belonging synchronically to the paradigm of "have" is certainly plausible, though not necessarily required (see also n.16).
16. I say this is an "apparent present form" since the paradigm is not identical with that of the present of "have", though, as mentioned in n.15, it does overlap considerably with the present

paradigm of the full verb. In any case, the auxiliary verb here is closer in form to the present of "have" than to the past, which has the root *av-* running throughout the paradigm (e.g., 1SG *aveam* "I had", 2PL *aveați*, etc.).

17. To be sure, as Harris points out, the French *passé composé* can also have the value of a true perfect, but the fact that it has replaced the *passé simple* is what is significant here by way of comparison with Daco-Romanian.
18. I say "drift" since the chronology of these changes in French, Spanish, and Italian — the Castilian changes are quite recent, for instance — means that they must be taken as independent but parallel innovations in these languages. Nonetheless, one can speculate that there is something in the set of structural oppositions in the verbal system inherited from Proto-Romance that might make such developments particularly natural, hence the product of a Romance "drift".
19. Interestingly, as Parkinson (1987:269) notes, there is a periphrastic past perfect in Portuguese, but it "is rarely used in colloquial registers", its synthetic counterpart being far more common. See also n.21 below.
20. Dialectally within present-day Daco-Romanian, and indeed also in earlier stages of Romanian, perfect forms with "be" as the auxiliary occur in parts of the system other than the future perfect, e.g., Transylvanian *am fost cântat* "I have sung" (literally: "I-have/AUX been sung"); see Rosetti 1968 and the very useful summary of verbal forms and dialectal perfects in Stefanescu (1997:277ff.).
21. So also Mallinson (1987:314); see however n.19 for a parallel in Portuguese to this Aromanian preference for the synthetic form, which could suggest an independent origin for this Aromanian development.
22. Also, the past perfect is more common in Greek than the present perfect, a fact which constitutes some of the evidence for taking the past perfect to be historically prior to the present perfect, as discussed by Joseph (1983, 1999b).
23. A further conclusion would be that the sort of contact Daco-Romanian speakers had with speakers of other Balkan languages, e.g., Bulgarian, was of a different character, perhaps in a different social setting (e.g., more focused on church-related matters than on day-to-day, side-by-side living).
24. Other studies can be cited. Rudin 1988, for example, investigates the properties of different strategies for the multiple fronting of *wh* elements cross-linguistically, and attempts to "integrate them into a general typology of *Wh*-movement and multiple-*Wh*-constructions" (p.488); Romanian figures prominently here, since Daco-Romanian and its Balkan neighbor Bulgarian provide the only examples in her survey of multiple *Wh*-Fronting languages with only syntactic *Wh*-Movement.
25. The transitivity part of this bipartite condition — i.e., both not transitive and not objectless — would cover the reflexive passive, under the assumption that it is a detransitivizing construction. However, if the Romanian reflexive passive in the complement of (14b) is analyzed so that the ostensible reflexive pronoun *se* is taken as a realization of the logical object of "do" (as is the case in some treatments of reflexive passives, as opposed to regular passives, in Relational Grammar), then the bipartite condition can be reduced to a single condition, namely "not objectless".
26. As Joseph (1980b; 1983:227) notes, Albanian also shows a finite complement clause in *tough*-Movement sentences with a resumptive pronoun corresponding to the moved nominal, thus providing another Balkan point of comparison here.

27. Bach & Horn (1976: 271–2), for instance, have cited as acceptable the following English sentence with “tough”-Movement out of a finite complement clause:
 - (i) Walter_i is hard for me to imagine that anyone would look at Ø_i.
28. For example, Grosu & Horvath 1987 see the constraint as subsumed under a more general principle that “complement clauses with an empty operator in their COMP binding a syntactic variable are non-finite”, though see Joseph (1978/1990: 197n.B) for some counter-discussion.
29. That is, whatever state results in a language from contact with speakers of another language is by definition a natural human language, one that must therefore be accommodated into Universal Grammar; the absence of a role for Universal Grammar in guiding the direction of borrowing is especially evident if, as Thomason & Kaufman 1988 have argued, there are no linguistic constraints on what can be borrowed, only social constraints.
30. Furthermore, the parallel noted here between Romanian and Greek is a rather abstract and “deep” one. It is thus of a sort that seems not to emerge from language contact, for contact-induced structural parallels generally depend on shared surface structures resulting from intense and sustained contact, often with some degree of bilingualism. Numerous examples of such surface-oriented parallels are evident in the Balkan *Sprachbund* (see (1c) above and the discussion in section 2), and the structural convergences found in Kupwar village in India, as described by Gumperz & Wilson 1971, are of the same type; see also Joseph 1999a for some discussion of the role of surface structures in contact situations.
31. The reduction of *îmi* to *mi* is regular in the presence of another weak pronoun, here *î*.
32. In Greek, as argued in Joseph (1978/1990, 1980b, 1983, and elsewhere), the decisive factor in determining the placement of weak object pronouns is finiteness; the patterns of placement are similar in Romanian but finiteness has a different role in the overall verbal system, inasmuch as in Romanian the category of infinitive (i.e., a verbal noun used in complementation) is more robust than in Greek (where it is absent altogether). Thus the precise characterization of the relevant triggering features for weak object placement may differ between Greek and Romanian but the surface patterns are quite similar.
33. Piedmontese Italian does show postverbal positioning of weak pronouns in compound tenses, so that the Romanian situation in general is not unparalleled in Romance; however, this postpositioning in Piedmontese is not restricted to just the feminine singular forms, so that Daco-Romanian is unique in the particular restriction it shows. See Parry 1995 for presentation and discussion of the Piedmontese facts. I am grateful to my colleague Dieter Wanner for clarification of these facts and for directing me to appropriate sources.

References

- Bach, Emmon & George Horn. 1976. “Remarks on ‘Conditions on Transformations’”. *Linguistic Inquiry* 7.265–299.
- Banfi, Emanuele. 1985. *Linguistica balcanica*. Bologna: Zanichelli.
- Boretzky, Norbert. 1968. “Der bestimmte Artikel nach Präpositionen im Albanischen”. *Zeitschrift für vergleichende Sprachforschung* 82.129–152.
- Comrie, Bernard, ed. 1987. *The World's Major Languages*. London: Croom Helm.
- Germano, Girolamo. 1622. *Grammaire et vocabulaire du grec vulgaire publiés d'après l'édition de 1622*. (ed. by Hubert Pernot, 1907). Paris.

- Green, John N. 1987. "Spanish". Comrie 1987. 236–259.
- Grosu, Alexander & Julia Horvath. 1987. "On Non-Finiteness in Extraction Constructions". *Natural Language and Linguistic Theory* 5.181–196.
- Gumperz, John & Robert Wilson. 1971. "Convergence and Creolization. A case from the Indo-Aryan/Dravidian border in India". *Pidginization and Creolization of Languages* ed. by Dell Hymes, 151–167. Cambridge: Cambridge University Press.
- Hamp, Eric P. 1989. "Yugoslavia — A crossroads of Sprachbünde". *Zeitschrift für Balkanologie* 25.44–7.
- Harris, Martin. 1987. "French". Comrie 1987. 210–235.
- Joseph, Brian D. 1978/1990. *Morphology and Universals in Syntactic Change: Evidence from Medieval and Modern Greek*. Ph.D. Dissertation, Harvard University; printed and distributed by Indiana University Linguistics Club [Expanded and updated version published (1990) in Outstanding Dissertations in Linguistics series, New York: Garland Publishing, Inc.].
- . 1980a. "A New Convergence involving the Balkan Loss of the Infinitive". *Indogermanische Forschungen* 85.176–187.
- . 1980b. "Linguistic Universals and Syntactic Change". *Language* 56.345–370.
- . 1983. *The Synchrony and Diachrony of the Balkan Infinitive. A study in areal, general, and historical linguistics*. (= *Cambridge Studies in Linguistics, Supplementary Volume*.) Cambridge: Cambridge University Press.
- . 1999a. "Is Balkan Comparative Syntax Possible?" To appear in *Balkan syntax in a Comparative Light* ed. by Maria-Luisa Rivero & Angela Ralli. Oxford: Oxford University Press.
- . 1999b. "Textual Authenticity: Evidence from medieval Greek". To appear in Susan Herring, Piet van Reenen & Lene Schoesler, eds., *Textual Parameters in Ancient Languages*. Amsterdam & Philadelphia: John Benjamins.
- Koerner, Konrad. 1982. "August Schleicher and Linguistic Science in the Second Half of the 19th C". *General Linguistics* 22.1–39 [reprinted in revised and expanded form in Koerner 1989: 327–375].
- . 1984. "Franz Bopp (1791–1867)". *Aschaffener Jahrbuch* 8.313–319 [reprinted in revised and expanded form in Koerner 1989: 291–302].
- . 1987. "Friedrich Schlegel and the Emergence of Historical-Comparative Grammar". *Lingua e Stile* 22:3.341–365 [reprinted in revised and expanded form in Koerner 1989: 269–290].
- . 1988. "Jacob Grimm's Place in the Foundation of Linguistics as a Science". *Word* 39:1.1–20 [reprinted in revised and expanded form in Koerner 1989: 303–323].
- . 1989. *Practicing Linguistic Historiography*. (= *Studies in the History of the Language Sciences*, 50.) Amsterdam & Philadelphia: John Benjamins.
- Mallinson, Graham. 1987. "Rumanian". Comrie 1987. 303–321.
- Newmark, Leonard, Philip Hubbard & Peter Prifti. 1982. *Standard Albanian. A reference grammar for students*. Stanford: Stanford University Press.
- Parkinson, Stephen. 1987. "Portuguese". Comrie 1987. 260–278.

- Parry, M. Mair. 1995. "Some Observations on the Syntax of Clitic Pronouns in Piedmontese". *Linguistic Theory and the Romance Languages*. (= *Current Issues in Linguistic Theory*, 122) ed. by John C. Smith & Martin Maiden, 133–160. Amsterdam: John Benjamins.
- Portius, Simon. 1638. *Grammatica Linguae Graecae Vulgaris* (ed. by Wilhelm Meyer, 1889). Paris: F. Vieweg.
- Reichenkron, Gunther. 1962. "Der Typus der Balkansprachen". *Zeitschrift für Balkanologie* 1.91–122.
- Rosetti, Al. 1968. *Istoria Limbii Române*. Bucharest: Editura pentru Literatura.
- Rudin, Catherine. 1988. "On Multiple Questions and Multiple WH Fronting". *Natural Language and Linguistic Theory* 6.445–501.
- Sandfeld, Kristian. 1930. *Linguistique balkanique. Problèmes et résultats*. Paris: Honoré Champion.
- Schaller, Helmut. 1975. *Die Balkansprachen. Eine Einführung in die Balkanphilologie*. Heidelberg: Carl Winter Universitätsverlag.
- Stefanescu, Ioana. 1997. *The Syntax of Agreement in Romanian*. Ph.D. Dissertation, City University of New York Graduate Center.
- Thomason, Sarah G. & Terrence Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.
- Thumb, Albert. 1912. *A Handbook of the Modern Greek Language. Grammar, Texts, Glossary*. Chicago: Argonaut, Inc. Publishers [reprinted 1964].
- Trask, Robert Lawrence. 1997. *The History of Basque*. London: Routledge.
- Vincent, Nigel. 1987. "Italian". Comrie 1987. 279–302.
- Wanner, Dieter. 1993. "Multiple Clitic Linearization Principles". *Linguistic Perspectives on the Romance Languages* (= *Current Issues in Linguistic Theory*, 103) ed. by William Ashby, Marianne Mithun, Giorgio Perissinotto & Eduardo Raposo, 281–302. Amsterdam: John Benjamins.

PART VII

Germanic, Caucasian, and Asian Linguistics

CHAPTER 37

Toward “a Complete Analysis of the Residues”

On regular vs. morpholexical approaches to OHG umlaut

David J. Holsinger

Joseph C. Salmons

University of Wisconsin–Madison University of Wisconsin–Madison

1. Introduction

Umlaut has been a central topic in historical linguistics since Grimm first described the basic patterns of these regressive vocalic assimilations in Germanic. This paper traces one chapter in the history of umlaut studies, with a focus on the chronic conflicts between the scope of the data set and the effort to make broad generalizations within the confines of regular sound change. That tension, especially in Old High German umlaut, runs from the Neogrammarians to the present day. After some basic background and data in §1, §2 describes progress in umlaut studies, driven by progress in theories of phonology and sound change. §3 outlines one contemporary continuation of that old problem — the role of “morpholexical conditioning” in early Old High German — and proposes a way of reconciling a major difference on this issue by careful review of both relevant data and theory. Aside from an original proposal on the relationship between regular sound change and morpholexical conditioning, the paper aims to provide what the recipient of this volume, Konrad Koerner (1995: 6), has called “the presentation of the linguistic past as an integral part of the discipline itself”.

2. Background

One basic friction in the study of sound change was neatly captured by Bloomfield (1933:354): "The real point at issue is the scope of the phonetic correspondence-classes and the significance of the residues. The neo-grammarians claimed that the results of study justified us ... in seeking a complete analysis of the residues". While most scholars accept this challenge in theory, umlaut presents such dense and variable data that few seem to entertain the possibility of phonetic conditioning across the full dataset in Old High German (OHG). Moreover, by Modern German, umlaut is — in the view of most — a morphological process, no longer strictly phonological.

At the center of the long debate over the historical unfolding of umlaut in Germanic are the OHG data. Traditional Neogrammarian works note two distinct umlaut processes taking place in OHG. An early process (illustrated with OHG data in (1) below), commonly called primary umlaut, raised *a* to *e* when followed by *i* or *j*, save when certain consonant clusters intervene. This very limited phonological change nonetheless has far-reaching effects, resulting in widespread alternations between *a* and *e* in many grammatical categories. Nonetheless, primary umlaut is the only one of the various umlaut processes which show themselves to be quite regular in the same environments in roughly contemporaneous Old English texts and which show up in other Germanic dialects (save Gothic, attested too early to reflect umlaut).

- (1) Examples of OHG Primary Umlaut: Verbal forms (Braune & Eggers 1986)

Present paradigms: *faru* "I drive" vs. *ferist* "you [sg.] drive"

Past Part. alternations: *walta* vs. *welita* "ruled, dominated" (< -*ta* vs. -*ita*)

Comparative Germanic: *nerien* "to save" vs. Gothic *nasjan*

By the end of the OHG period, however, the linguistically savvy Notker der Deutsche, translator of the works of Boethius, begins to indicate the collapse of the old diphthong *iu* with an original *u*: in umlaut positions, both presumably [y:]. These mark the first systematic orthographic appearances of 'secondary umlaut', a more inclusive set of mutations which affects all non-front vowels in the same environment as primary umlaut, namely, before *i* ~ *j*, as illustrated briefly in (2). The limitations of the Latin alphabet in expressing front rounded vowels may have retarded the appearance of this process somewhat, but the process must have been completed as a phonological process by the Middle High German period, when the triggers were reduced to schwa.

(2) Non-primary umlaut

<u>OHG</u>	<u>MHG</u>	<u>Gloss</u>
<i>gruoni</i>	<i>grüene</i>	"green"
<i>skoni</i>	<i>schæn(e)</i>	"beautiful"

Nonetheless, even Middle High German manuscripts do not regularly indicate umlaut. Scholars of Germanic historical phonology thus find themselves faced with the question of whether secondary umlaut may have, in fact, been present during the OHG period. Despite the failure of any High German manuscript prior to the late 10th C. to indicate secondary umlaut systematically, the process was not only completed but indicated orthographically in Old English manuscripts. This presence of the same types of vocalic mutation in the same environments in two closely related languages has led many scholars to posit umlaut as a characteristic of the proto-language rather than as an independent, parallel development in both languages, e.g., Antonsen and Penzl, as discussed briefly below.

The central question is how to interpret the comparative historical record. Do the OHG data present us with indirect or direct evidence of regular sound change, or do they represent the effects of a long-past change obscured by massive analogy and morpholexical reanalysis, only belatedly and begrudgingly tolerated by an inadequate orthography? In this paper, we trace the pursuit of regularities found and generalizations proposed about OHG umlaut and look at the dwindling set of forms attributable to analogy and morpholexical conditioning.

3. A brief history of OHG umlaut studies

One recent work (Krygier 1997: 117) sees the history of *i*-umlaut studies as a prime example of "theory recycling", driven by a lack of real theoretical progress: "When there is no new framework for data analysis, the temptation to resurrect an old hypothesis and reformulate it in the linguistic jargon of the day is particularly strong". Krygier concludes that much current work on umlaut reiterates old views without apparent progress. His pessimism, however, comes from a focus on elusive motivations for umlaut, whereas we focus on its unfolding, an area where tangible gains have been made in this century on empirical foundations carefully laid by early historical-comparative linguists. Those same scholars, of course, charted the rocky waters between regular sound change and analogy, alluded to above.

W. Freeman Twaddell's 1938 "A Note on Old High German Umlaut" was the first work to apply modern phonological theory to the problem. His concise account remains the classic statement, and by providing a model of how sound change becomes phonemic, it represents a landmark for phonemic theory. Twaddell captures the critical difference between conditioned allophonic variation and the eventual phonemic distinctions resulting from the loss of triggers, but this teleological orientation forces an understanding of sound change as catastrophic (very unlike Neogrammarian views, cf. Collinge 1995: 205). The abstractness of his model, by which only the phonemic status of a given element is relevant, provides insight into the ends but not the means of sound change. This leads to a reduction of the data set analyzed, since the blocking environments (for example, clusters of *ht* and *hs*, which block primary but not secondary umlaut) can only be described as exceptions rather than as part of a unified process.

Later structuralists, notably Penzl (1949) and Antonsen (1969), apply the phonemic model to the full range of umlaut data, and project the full set of vocalic mutations back to the proto-language. Here we see the full extension and implication of the teleological orientation apparent in Twaddell's model: Rather than admit the possibility of linguistic drift — the independent operation of equivalent sound changes in related languages — structuralists proposed that allophonic variation in umlaut environments was already present in the proto-language and simply unfolded at different rates as the conditioning factors were lost in each of the daughter languages. Without a model of the PROCESS of sound change, this is the only plausible view of umlaut in Germanic, but nonetheless the logical extensions of this model doom the analysis. Especially in the coastal Dutch dialects showing umlaut only of short vowels, the explanation of exceptions to umlaut by means of analogy is simply untenable — many Dutch forms with umlaut of short *a* are without any reasonable basis for analogy, such as *bed* from WGmc. **badi*. Appeal to analogical change here denies patently regular sound change in order to maintain the view that any similar pattern among related languages is inherited. The model fails primarily due to its abstract view of sound change: It succeeds in capturing the core regularity of the processes, but excludes variability and phonological exceptions from the data.

The advent of generative phonology led quickly to a reanalysis of the OHG umlaut data, focusing not on the means by which allophonic variation became phonemic, but rather on describing the umlaut process. King (1969) marks a radical departure from structuralist analyses, formalizing primary and secondary umlaut not just as chronologically distinct but also phonologically independent processes, shown in (3). The disparate historical shifts involved in umlaut appear

as the orderly application of simple, well-defined phonological rules to the same set of inputs. Rather than abstracting away from the process, early generative work concerned itself with the interrelation of sound changes as the application of a set of rules to a single form.

- (3) Early generative umlaut rule (after King)

$$\left[\begin{array}{c} V \\ <-long> \end{array} \right] \rightarrow \left[\begin{array}{c} -back \\ <-low> \end{array} \right] / \text{---} C_0 \left[\begin{array}{c} -cons \\ -back \\ +high \end{array} \right]$$

It is here, however, that generative phonology no longer provides a convincing model of the outcome of sound change. By proposing that the same forms continued to serve as the basis for a sequence of sound changes, generative rules leave us without a resolution to the question of when umlaut ceased to be a strictly phonological process derivable from one set of inputs. This same question plagues much modern umlaut research: Umlaut, even in Modern German, has the look in some ways of a simple and straightforward phonological rule, but at what point and to what extent does the rule have to be learned independently of phonology, either as a morphological rule or as part of a lexical representation? By simplifying the formal representation of umlaut, generative analyses have introduced a new degree of complexity, attributable only to the understanding of grammar that generative phonology provides: A derivation by rules is not necessarily simpler than a set of distinct lexical forms clustered around a morphological pattern. The simplicity of phonological rules comes at the expense of clear, direct relations among the lexicon, morphological patterns, and phonology, issues to which we now turn.

4. Current appeals to morpholexical conditioning

In the remainder of this paper, we focus on two contemporary directions in umlaut research, a split driven by the same desire to capture regularity in a rigorous way on the one hand and a desire to accommodate all wrinkles in the data on the other. An approach driven by the latter is embodied by Janda (1998:169), who sees "the forces of phonetics and phonology versus those of morphology and the lexicon" as "combatants" with the latter having "carried the day". This work, represented prominently by a string of papers culminating with Janda (1998) and built on earlier work by Orrin Robinson (1975, 1980), maintains that umlaut was morphologized by early OHG. Umlaut was used in certain

morphological categories and banned in others, but it had not, in this analysis, been phonologically conditioned for centuries.¹ The other approach, dubbed Ingenerate Umlaut by Iverson & Salmons 1999, stresses phonological regularities in the unfolding of umlaut across Germanic, such as the blocking environments for primary umlaut and the cline of umlautlessness across Netherlandic dialects. This school of thought (including, broadly speaking, work by Buccini, Davis, Howell, Iverson, Salmons, and Smith) focuses on the inherent, co-articulatory basis for umlaut and therewith on finding phonological regularities where earlier structuralist and early generative approaches resorted to analogical accounts. For example, the blocking environments of liquid or /x/ plus obstruent for OHG umlaut (cf. *nahti* "night") and the "umlautless residues" of Upper German (southern dialectal *muck* vs. Standard German *Mücke*, etc.) can be explained in terms of feature geometry and coarticulation rather than analogy, which is seen as becoming a major factor only later, in late OHG and into MHG.

This debate is not about whether Modern German umlaut is morphologically conditioned but about when and how morphological umlaut developed from an originally phonological process. The critical juncture here comes in the class i weak or *-jan* verbs, so-called *Rückumlaut* verbs, which fail to umlaut in their preterit subjunctive forms, illustrated in (4). Morpholexicalists take exactly this bit of data as the linchpin for arguments denying the phonological regularity of umlaut already in OHG. Those seeking regularity claim that even such forms as these (at least as found in early texts) should yield to careful phonological analysis: Orthographic variability aside, the full phonological conditioning of OHG umlaut should be apparent even beneath the effects of later analogical restructuring. Indeed, pret. subj. forms are the best-known set of data for which proponents of ingenerate umlaut have not offered a phonological account. In this section, we seek a resolution of this contentious issue.

(4) Class i weak or *-jan* verbs in OHG (Braune & Eggers 1986)

<u>inf.</u>	<u>pret.</u>	<u>pret. subj.</u>	<u>past part.</u>	<u>gloss</u>
<i>brennen</i>	<i>branta</i>	<i>branti</i>	<i>gibrennit</i>	"to burn"
<i>sterken</i>	<i>starcta</i>	<i>starcti</i>	<i>gistarkit</i>	"to strengthen"
<i>zellen</i>	<i>zalta, zelita</i>	<i>zalti, zeliti</i>	<i>gizalt, gizelit</i>	"to say, tell"

What is the relationship between phonological and morpholexical accounts of OHG umlaut? The morpholexical view rests on the absence of phonological conditioning. Robinson works hard to find a phonological basis for the lack of umlaut and resorts to morpholexical conditioning when that fails.² Janda, in contrast, complains of the "overwhelmingly phonological bias" (1998:172) of umlaut studies, and he largely discards phonology from the start. We side with

Robinson in believing phonologically regular accounts to be preferable to analogical or morpholexical accounts: First, a phonological account, if available, provides rigor, and in some sense explains the attested patterns better than irregular analogy.³ Second, it is clear that umlaut, at some past point, had its roots in phonology, and slowly evolved into a more and more morphological phenomenon. The earliest history of umlaut must be sought in the sound system; to abandon such a search is to leave one of the most important chapters in Germanic linguistic history untouched. Finally, we must consider the price paid for a morpholexical account, which has problems with chronology (see footnote 1, above), and must posit vast patterns of analogy with suspiciously phonological-looking conditioning. One may, in short, conclude that a morpholexical approach has no trouble at all with the data, but that is because a morpholexical account sidesteps the earliest (pre-)history of umlaut and provides a less principled account of synchronic OHG grammar. At any rate, both phonology and morphology play significant roles in the history of umlaut and the question is more WHEN than WHETHER each is at play. The morpholexical position suggests or argues that we cannot see back to the phonological process; the ingenerate umlaut view insists that we must look thoroughly.

The starting point for morpholexical approaches is Wilmanns' claim (1897: 252, cited by Robinson 1980 and Janda [1998:176]) that these forms show "nowhere a trace of umlaut" ("nirgends eine Spur des Umlauts"). Robinson describes the pret. subj. forms as showing "total resistance to umlaut" as a category (1980: 450). This set provides an exception to umlaut that "stands out from all the others ... by being exceptionless even in the oldest available texts" (1980: 450). Janda (1998:178) picks this claim up, e.g., "OHG indeed had an entire morphological category ... in which — for completely non-phonological, morpholexical reasons — umlaut simply did not apply". If this data cannot be reconciled with regular sound change, we, like Janda, will be forced to accept that umlaut was morpholexical from early OHG. The other alternative would be to treat these data as exceptional. Either way, the phonological (pre-) history of umlaut would have to be seen as unreconstructable.

5. Fresh data and another approach to pret. subj.

To avoid that sad state of affairs, we first present some old but overlooked data on this question and then integrate that data into a broader view of umlaut. Janda (1998:175) holds up Raven's 1963 compendium on *-jan* verbs as an indication of just how common umlautless *Rückumlaut* forms were. If one actually peruses

the data in Raven, the number of relevant forms shows itself as small, hardly "a very large group of OHG words". While the *-jan* class was indeed large, only a fraction of the verbs are of the *Rückumlaut* type (those showing preterit vocalism of /a/) and pret. subj. forms are quite ill-attested in the verb system. The verb *senten* "to send", source of the often-cited pret. subj. *santa*, is attested in OHG over 200 times (counting from Raven), but with only three pret. subj. forms. To take a non-*Rückumlaut* example, *furhten* "to fear" is attested more often than *senten* but lacks any pret. subj. attestations.

Even within that limited set of attested umlautable pret. subj. forms in OHG, a number of umlauted examples are in fact attested. Let us turn first to a small set of forms that Janda himself treats in a footnote, namely *zeliti(n)* from *zel(l)en*. These forms maintain an original linking vowel lost early in most long stem *-jan* verbs. Janda argues that it was this *Bindevokal*, not the suffix vowel which triggers umlaut and that, in at least one text, all forms of *zel(l)en* in the preterit and pret. subj. show *e* rather than *a* vocalism, concluding that these forms were lexicalized. The facts are, however, more complex. First, Otfried uses umlauted and umlautless forms in the pret. and pret. subj. of *zel(l)en*, as shown below in (5), with sources indicated.⁴

(5) Examples of *zel(l)en* in Otfried

	Umlautless	Umlauted
Pret.	<i>zalta</i> 2, 6, 17	<i>zelita</i> 2, 7, 9
Pret. subj.	<i>zalti</i> 1, 11, 5	<i>zeliti</i> 2, 7, 42

While medial *-i-* does indeed correlate with umlaut here (we count nine forms in Raven with *-i-* and umlaut for this verb), and thus supports Janda, the kind of variation shown in (5) above undermines the case for lexicalization. The assumption of umlaut failure leads to the hypothesis of morpholexical conditioning, and thus evidence for phonological conditioning must be dismissed as lexical rather than admit that the premise was false. Nonetheless, Janda claims that these forms do not disconfirm "the generalization that there is no OHG text in which past subj.-marking /i/ itself ever triggers umlaut" (1998:214). (6) gives a broader set of forms retaining medial *-i-*, none of which seems likely to be lexicalized.

(6) Other *-jan* verbs with retention of medial *-i-* and umlauted pret. subj.

<i>cheletin, queliti</i> , etc.	Otfried, Glosses, etc.	< <i>chel(l)en</i> "to torture, torment"
<i>kisezzatin</i>	Glosses	< <i>gi-sezzen</i> "to set" (vs. <i>gisazti</i>)
<i>hebiti</i>	Glosses	< <i>haben</i> "have" (vs. <i>habeti</i> , etc.)
<i>gisegiti</i>	Glosses (2x)	< <i>seg(g)en</i> "say" (vs. <i>sageti</i> , etc.) ⁵

This set, like *zeliti*, reflects retention of the old medial vowel, although some forms, like *cheletin*, show reduction of that vowel rather than the presence of an unambiguous umlaut trigger. Across the OHG corpus, all these verbs attest variation between umlauted and non-umlauted pret. subj. forms, e.g., Otfrid's *qualtin* "they would have tormented". Already a modest and familiar phonological generalization can be made: Where the *Bindevokal -i-* is retained, umlaut tends to occur. If the pret. subj. were as a morphological category somehow resistant to umlaut, this would not be true.

In a morpholexical view, the very absence of umlaut marks the pret. subj. as distinct from the preterit indicative. The forms in (7), again culled from a survey of Raven with a contextualized example in (8), provide yet more direct and powerful counterevidence. That is, these examples are directly at odds with a morpholexical analysis.

- (7) Preterit Subjunctive *-jan* verbs with primary umlaut but without *-i-*
- | | | |
|------------------|----------------|----------------------------------|
| <i>uuenti</i> | Otfrid (see 8) | < <i>uuenten</i> "to turn", etc. |
| <i>prehten</i> | Glosses | < <i>bringen</i> "to bring" |
| <i>zebrehten</i> | Glosses | < <i>zuo-bringen</i> |

- (8) *fārawa ni uuenti, fol bistu gōtes ensti* (1, 5, 18)

The first form shows umlaut in a typically non-blocking environment, with *-nt-* intervening between trigger and target.⁶ The second and third exemplify the strongest and most dialectally widespread blocking configuration. Indeed, the vowel in *brahti* was presumably long, indicating that this would be an instance of secondary, not primary, umlaut. In a strictly morpholexicalist view, this becomes yet more difficult to explain, since pret. subj. umlaut is attested beyond even primary umlaut. And, once again, across OHG these lexical items show umlauted and umlautless forms, making lexicalization improbable. More importantly, after appealing to the absolutely regular absence of umlaut to justify a morpholexical approach, it becomes awkward to appeal back to morphological factors to justify the unexpected presence of umlaut alternations in these verbs.

"Why aren't there any relics?", asks Robinson (1980: 451). In fact, even in the strictest sense of 'relic', there are a few, as shown in (7) and (8) above. The notable TOTAL resistance to umlaut is illusory; this ENTIRE morphological category is not immune to umlaut. Are we left with nothing more than a curious, perhaps statistically indeterminate, tendency to employ umlautless variants in this morphological category?⁷

Let us propose a step beyond that cynical assessment, one which may bring the two positions closer than they have previously been. Robinson's greatest

insight is that OHG shows the prehistory of umlaut's morphologization, just as the masculine *n*-stems show the first signs of eliminating umlaut in the singular within the OHG period. This is a particularly plausible account for pret. subj. forms in the history of OHG. As noted, this category, contra Janda, is small and ill-attested, thus a natural place for early lexical oddities to creep in. Moreover, traditional accounts often and rightly connect the absence of umlaut in pret. subj. forms to blocking environments: Given that verb stems of this class end in consonants or clusters, often fricatives and liquids, suffixation of *-ti* yields a striking percentage of forms that should block umlaut: *ahtin*, *brahti*, *falti*, *gahti*, *martin*, *gistarhti*, etc. In fact, the only really common pret. subj. verb *-jan* forms without *-i-* are those from *bringen* (*brahti*, etc.) with 24 occurrences and *denchen* (*dahti*, etc.) with 22.⁸ Among those with medial *-i-*, *zel(l)en* has the most pret. subj. forms, with the alternations between two types sketched above. In counting pret. subj. forms from Raven in the entries through the letter G, 50 were classic blocking environments (48 *-ht-* and 2 *-lt-*), 4 more were trisyllabic, and 11 were typically non-blocking environments (but including 3 tri-consonantal clusters). The presence of an additional medial vowel in other forms further dilutes the amount of umlaut expected, as may complex tri-consonantal clusters. A large percentage of all remaining *-jan* pret. subj. forms are of the *zeliti* type, with medial trigger and umlaut. For these reasons, of all classes in OHG verbal morphology, this may be the most naturally resistant to umlaut. Whenever morpholexical conditioning of umlaut may have arisen, the phonetic shape of pret. subj. forms makes them a logical starting point for it. The riddle is reduced to a very small set of forms like *krati* 'to crow'.

Janda (1998 and elsewhere) translates Robinson's earlier work into Lexical Phonology and pushes the dating back farther, arguing that umlaut was morpholexical throughout OHG, and that all phonological conditioning had disappeared in pre-OHG. Janda's arguments, however, rest on a questionable interpretation of the apparent morphological patterning of umlaut in OHG. Even if most pret. subj. *Rückumlaut* verbs showed his proposed morphological umlaut blocking, the exceptions (forms *with* expected umlaut) look suspiciously like they reflect a lexical rule. Janda frequently invokes Lexical Phonology in defence of his own position, but he neglects the phonological basis for what eventually become morphological rules. Instead of looking for the basis by which exceptional forms were marked in either the lexicon or in morphological derivation, Janda sees the end result of the shift to morphological umlaut as a stable pattern throughout all of OHG. As noted earlier, umlaut blocking in pret. subj. forms such as *falti* or *ahtin* clearly has a phonological basis: The clusters *-ht-*, *-lt-*, *-rt-*, derived in pret. subj. forms by affixation of *-ti*, are blocking environments for primary umlaut.

It is no stretch to propose that the high frequency of phonological umlaut-blocking in *Rückumlaut* verbs comes to be reanalyzed as a set of exceptions to the phonological umlaut rule. Such conditions on the application of a phonological rule reflect the Elsewhere Condition: A more specific rule applies before a more general one. The pret. subj. suffix *-ti* is marked as non-umlauting, the exceptional case, while other *i*-ful suffixes (including the allomorph *-i-ti* of the pret. subj. affix) continue to cause umlaut.

In Kiparsky's view (1988:373), "morphological conditions on sound change originate as variable constraints at the postlexical stage; they become conditions on rules only when the rules enter the lexical phonology". We suggest here simply that primary umlaut became a rule of lexical phonology early in OHG. Rather than serving as evidence for any well-established morphological basis for umlaut, the pret. subj. forms reflect the lexical status of primary umlaut, as proposed by Iverson & Salmons (1996). Once umlaut blocking is extended as a lexical phonological rule in a class of forms, the propagation of both umlaut and umlaut blocking into more clearly morphological functions follows the model of the original lexical phonological rule. Robinson (1980) proposes an account of umlaut failure based on the morphological functionality of umlaut in the pret. subj. as opposed to other categories: In *Rückumlaut* verbs, if the pret. subj. shows umlaut, it resembles a functionally different category — the present tense — while in other verbs, umlaut functions as a salient marker of the category "subjunctive".⁹ Such patterns are familiar from the modern literature on lexical or 'variable' rules. For example, Labov (1981:286, see also Kiparsky 1995:650) shows the absence of Philadelphia /æ/ tensing (a clearly lexical rule in this dialect, while postlexical in some other American dialects) in past forms of strong verbs ending in nasals, a non-change with close parallels in the history of English class III strong verbs.

The pret. subj. verbs, as we can agree with Robinson, represent the clearest case of morphophonological conditioning for umlaut in OHG, but as the beginnings of the morphologization of umlaut, still a clearly lexical phonological process. The rarity of such forms outside of blocking environments and forms with medial *-i-* makes the chronology of this process within OHG difficult, if not impossible, to pin down.

In this paper, we have concentrated on early OHG data and a word about the broader context is in order here. Following one of the main lines of argument for Ingenerate Umlaut (especially Iverson & Salmons 1996), we see primary umlaut as both chronologically prior to non-primary umlaut and as far more directly and clearly phonetically-phonologically conditioned, thus our focus on the Old High German period. Middle High German shows a pattern of *Rück-*

umlaut beyond verbs with short /a/, including indeed the verb “to fear” with blocking in pret. subj. forms (cf. Paul, Wichl & Grosse 1989: 66–67, 260, etc.), a point called to our attention by Rob Robinson (by electronic communication). Only from the earliest data and various umlautless “residues” can we catch sight of the purely phonetic-phonological character of primary umlaut. As the more general umlaut process unfolds in late OHG and into early MHG, it shows much exceptionality, including morphological sensitivity, something pointed out in earlier ingenerate analyses. In short, like Robinson, we see the umlautlessness of pret. subj. forms in MHG as an indication that umlaut is no longer a simple phonological rule.

6. Conclusion

We have argued a number of points in this paper. First, there is the mundane fact that the morpholexical interpretation of OHG umlaut rests on incomplete data, which alone calls into question the strong version of a morpholexical analysis. This allows a resolution of the ingenerate vs. morpholexical positions because umlaut during the OHG period comes to show the characteristics typical of lexical rules. Of greater relevance for the title and content of this volume, we hope to have shown the ongoing relevance and continuing contributions made possible by the Neogrammarian insistence on careful sifting of the data and Bloomfield’s “complete analysis of the residues”. In short, structural linguistics still must hold to those underlying principles.

The morpholexicalist focus on the pret. subj. *-jan* verbs has been justified, not because they provide a counterexample to phonological regularity, but because they allow a glimpse into the first stirrings of umlaut’s long journey toward morphologization. We suggest that the data provide us a remarkably precise picture of the unfolding of OHG umlaut. Howell & Salmons (1997) and Iverson & Salmons (1999) have made similar arguments for the phonological (pre-)history of umlaut, to which the present paper offers a morphological parallel.

Acknowledgments

We would like to thank the following for comments on this paper and/or discussions on this subject: Paul Houseman, Rob Howell, Greg Iverson, Monica Macaulay, and especially Rob Robinson. All the usual disclaimers apply.

Authors' address

David J. Holsinger & Joseph C. Salmons

Department of German

818 Van Hise Hall

1220 Linden Dr.

University of Wisconsin

Madison, WI 53706, USA

(D.J.H.) djholsin@students.wisc.edu

(J.C.S.) jsalmons@facstaff.wisc.edu

Notes

1. While Janda has not (yet) committed himself in print to a date for umlaut ceasing to be phonologically conditioned, in an oral presentation at the Workshop on Comparative Linguistics (Wayne State University, November 1997), he speculated that this would have been before 600 of the Common Era, so that we have no records of German when umlaut would have been phonologically active. Even aside from the traditional arguments dating umlaut later, problems arise with Otfrid's umlaut-triggering enclitic pronouns *drenkih* "I drank" for **drank ih* or *werfiz* "threw it" for **warf iz*. These must be treated as lexicalized in a non-phonological account, although positing lexicalization of "threw it" strikes us as at best ad hoc.
2. Robinson finds previous phonological accounts inadequate, writing for example that "we are ultimately thrown back to a purely analogical cause for the lack of umlaut in our category" (1980: 455). See Iverson & Salmons 1999 for additional discussion of the advantages of a phonologically regular account.
3. Analogy has often been the garbage can into which forms and sets of forms are thrown if they do not yield to an account in terms of phonologically regular sound change; while analogy certainly plays a role in linguistic change, it is all too common to write off all but the most straightforward sound changes as the results of analogical change. That was clearly the case in late structuralist approaches, where all blocking environments, umlautless residues, and much of Netherlandic umlaut had to be treated as analogical.
4. The better case for lexicalization of old *Bindevokal* would be *leg(g)en* (cf. Gothic *lagjan*), which, unlike the verb under discussion, shows almost uniform retention of *-i-* throughout the entire paradigm. Every form in *Raven* is umlauted and only one (*legta*) lacks a medial vowel.
5. The last two verbs in this class are originally weak class iii, *habên* and *sagên*, which in some dialects shifted to show class i forms in 2/3 sg. pres. and throughout the preterit. Rob Howell points out that this reverses the alternation on which morpholexical accounts of umlautless pret. subjs. rest, realignment of two extremely high-frequency verbs to yield umlauted preterit forms. This precisely counters the "pleasant redundancy" Robinson (1980: 457) sees as motivating loss of umlaut. As Howell further suggests, in this set of verbs, the umlauted preterits could be taken as the starting point for the analogy.

6. Nonetheless, homorganic nasal + obstruent sequences do appear to tendentially block umlaut, cf. Paul, Wiehl & Grosse (1989: 66–67, in part with reference particularly to pret. subj.) and Howell & Salmons (1997).
7. The remaining option for a morpholexicalist would be to argue that umlaut is statistically less common in pret. subj. forms. Sheer rarity of relevant pret. subj. forms would render statistical work difficult if not lacking in significance and, at any rate, no one has yet undertaken the challenge of exhaustive stochastic analysis of occurrence/non-occurrence of umlaut in OHG.
8. Recall that *brahti* (and *dahti*) probably had long rather than short /a/. As noted above, verbs of this type are attested with umlaut in OHG pret. subj. forms, which becomes yet more surprising if the vowel was long.
9. Greg Iverson points out that we still mark subjunctives with 'unexpected' verb forms today: "If I were", "Be it ever so humble", etc.

References

- Antonsen, Elmer H. 1969. "Zur Umlaufteindlichkeit des Oberdeutschen". *Zeitschrift für Dialektologie und Linguistik* 36.201–207.
- Bloomfield, Leonard. 1933. *Language*. New York: Holt, Rinehart & Winston.
- Braune, W. & H. Eggers. 1987. *Althochdeutsche Grammatik*. 14th ed. Tübingen: Niemeyer. [1st ed. Braune 1886.]
- Buccini, Anthony F. 1992. *The Development of Umlaut and the Dialectal Position of Dutch in Germanic*. Doctoral dissertation, Cornell University.
- Collinge, Neville. 1995. "The History of Historical Linguistics". *Concise History of the Language Sciences: From the Sumerians to the cognitivists* ed. by E. F. K. Koerner & R. E. Asher, 203–212. Oxford: Pergamon.
- Howell, Robert B. & Joseph C. Salmons. 1997. "Umlautless Residues in Germanic". *American Journal of Germanic Linguistics & Literatures* 9:1.83–111.
- Iverson, Gregory K., Garry W. Davis & Joseph C. Salmons. 1994. "Umlaut Blocking Environments in OHG". *Folia Linguistica Historica* 15.131–148.
- & Joseph C. Salmons. 1996. "The Primacy of Primary Umlaut". *Beiträge zur Geschichte der deutschen Sprache und Literatur (PBB)* 118.69–86.
- & ———. 1999. "Umlaut as Regular Sound Change: The phonetic basis of 'ingenerate umlaut'". *Language Change and Typological Variation: In honor of W. P. Lehmann* ed. by Edgar C. Polomé & Carol Justus, 207–224. (Journal of Indo-European Studies, 30).
- Janda, Richard D. 1998. "German Umlaut: Morpholexical all the way down from OHG". *Rivista di linguistica* 10.165–234.
- King, Robert D. 1969. *Historical Linguistics and Generative Grammar*. Englewood Cliffs, NJ: Prentice-Hall.
- Kiparsky, Paul. 1988. "Phonological Change". *Linguistics: The Cambridge Survey, vol. 1, Linguistic Theory: Foundations* ed. by Frederick J. Newmeyer, 363–413. Cambridge: Cambridge University Press.

- . 1995. "The Phonological Basis of Sound Change". *The Handbook of Phonological Theory* ed. by John Goldsmith, 640–670. Oxford: Blackwell.
- Koerner, E. F. Konrad. 1995. "History of Linguistics: The field". *Concise History of the Language Sciences: From the Sumerians to the cognitivists* ed. by E. F. K. Koerner & R. E. Asher, 1–7. Oxford: Pergamon.
- Krygier, Marcin. 1997. "Theory Recycling: The case of *i*-umlaut". *Studia Anglica Posnaniensia* 31.118–121.
- Labov, William. 1981. "Resolving the Neogrammarian Controversy". *Language* 57:2. 267–308.
- Paul, Hermann, Peter Wiehl & Siegfried Grosse. 1989. *Mittelhochdeutsche Grammatik*. Tübingen: Max Niemeyer. 23rd ed. [1st ed. Paul, 1881.]
- Penzl, Herbert. 1949. "Umlaut and Secondary Umlaut in Old High German". *Language* 25.223–240.
- Raven, Frijthof A. 1963. *Die schwachen Verben des Althochdeutschen. Band I: lang, mehr- und kurzsilbige -jan Verba*. Gießen: Schmitz. (Repr. University of Alabama Press, 1964.)
- Robinson, Orrin W. 1975. "Abstract Phonology and the History of Umlaut". *Lingua* 37.1–30.
- . 1980. "An Exception to Old High German Umlaut". *American Indian and Indoeuropean Studies: Papers in Honor of Madison S. Beeler* ed. by K. Klar, M. Langdon & S. Silver, 449–460. The Hague: Mouton.
- Smith, Regina. Forthcoming. "Testing Compensatory Theories of Umlaut". *Leuvense Bijdragen*.
- Twaddell, W. Freeman. 1938. "A Note on Old High German Umlaut". *Monatshefte* 30.177–181.
- Wilmanns, Wilhelm. 1897. *Deutsche Grammatik: Gotisch, Alt-, Mittel- und Neuhochdeutsch. Lautlehre*. 2nd ed. Strassburg: Trübner.

CHAPTER 38

Narrative Cohesion in the Kensington Runic Text

Robert A. Hall, Jr.†
Cornell University

1. Introduction

By now, late 1990s, it has been demonstrated beyond any reasonable possibility of doubt that the famous/notorious Kensington rune-stone is not a 19th C. fake.¹ On the contrary, no matter what the stone's vicissitudes may have been from 1362 A.D. down to its discovery by the Swedish settler Olof Öhman in November, 1898, the year 1362 fits well enough with what is known concerning both the fourteenth and the 19th C. backgrounds² so that we should regard the century-long debates over the genuineness of the inscription as closed.³

It must be emphasized that the dialect of the Kensington inscription is by no means identical to that of "the modern dialect of the Kensington region" (as alleged by Wahlgren and others). In the first place, there is no such thing as "THE modern dialect of the region", whether Swedish or Norwegian. Each settler family used its own native Scandinavian dialect, and no *koiné* was ever formed there. (I have checked this point with two outstanding specialists on Swedish and Norwegian dialects in the upper Mid-West).

Furthermore, the inscription has many features which are specifically mediaeval and are absent from modern Swedish and Norwegian (listed in both Hall 1994 and Nilsestuen 1994). The dialect of the inscription has been shown (especially by Nielsen forthcoming a, b) to be Old Bohuslänsk (intermediate between Norwegian, Swedish, and Danish). We must now proceed to more positive and productive studies of the structure and content of the text. These will, in their turn, furnish us with further corroboration of the authenticity of the Kensington stone.

For any discourse, short or long, to be effective in communication, it must

have a clear message, whose structure must "hang together", must be as coherent as possible.⁴ This cohesion can be attained by the use of whatever features of linguistic structure are available to the sender of the message — lexical, morpho-syntactic, phonological (including prosodic⁵). The Kensington text "hangs together" very tightly indeed, with remarkable compactness and concomitant brevity and clarity. To attain this result, the author or authors⁶ made use of stylistic devices⁷ on all of these levels.

The function of the Kensington text is to tell the story of a disaster which befell a group of Scandinavian explorers in what is now Minnesota, and to place it in its temporal and geographic context. One overall significant feature of the text is that its entire lexicon is Germanic, except for the element *AVM* (rr. 148–150), which almost all commentators agree in interpreting as "A[ve] V[irgo] M[aria]".⁸ The use of a totally Germanic lexicon would have made for immediate comprehension on the part of any possible Scandinavian readers.

The text falls quite naturally into three sections. The first of these (rr. 1–49) is, as it were, a prologue which sets the scene for the story of the main action (rr. 50–163), and the final section, on the side of the stone (rr. 164–222), gives further information concerning related matters. We shall take up each section in turn, discussing the specific features of linguistic structure that serve as bases for the stylistic devices used in the discourse.

2. The background

This is given in the first three lines, in which we are told of the European origin of the men, the purpose of their voyage, and its general direction in relation to Vinland, as narrated in rr. 1–49:

8 göter ok 22 norrmen þo [e-?] [þe-?] no opþagelsefarþ fro winlanþ of west
 "8 Goths⁹ and 22 Norsemen on [a?] [this?] voyage of discovery (or, possibly land-taking¹⁰) far to the west from Vinland".

In this section there is no verb, and the entire section constitutes, not a separate but incomplete sentence, but a long phrase in apposition to the pronoun *wi* (rr. 50–51) "we", which is the subject of the immediately following verb *habe* (rr. 52–55) "had". Such a construction is termed, in traditional grammatical analysis, a *proleptic apposition*, since it occurs before the element to which it is syntactically attached, not in dependency but in apposition.

The placing of an element of this type in advance of the beginning of the clause is known as "foregrounding" or "topicalization". The latter term is used

especially when the element in question brings out not only one, but several elements into relief as topics on which the rest of the clause will furnish comments — in this instance, the geographical region, the number of men in the group, the purpose of the voyage, and its geographical location and direction.

This type of construction is found in many conservative Indo-European languages, including Latin, Greek, Spanish, and other Romance languages, and Germanic languages such as Gothic and Old Norse (cf. Hall [1950] with many examples). Even more strongly than in that article, I am now persuaded that it is not a Latinism in the Germanic languages, but goes back to a common inheritance in Proto-Indo-European syntax. The absence of a verb in this long phrase has misled many commentators, from Holand on, to assume the ellipsis of a pronoun and verb **wi er* “we are” that would have preceded the present beginning of the sentence. This misunderstanding has been perpetuated by continual citation (by title only!) of Nygaard (1894), which has “elliptical” sentences as its main concern.¹¹

There are two main reasons, inherent in the text itself, why a hypothetical **wi er* “we are” should not be assumed. Both semantically and syntactically, its inclusion would weaken the effect of the beginning. The most important information to be communicated at the outset is not the simple identification of a group, large or small, referred to only as “we”, but the much greater list of details referred to in rr. 1–49. The semantic weakness of **wi er* would have been evident in speech as well, since it would have been relatively unstressed, in contrast to the full stresses which must certainly have been used with the words of the appositive element (cf. below, s.v. Phonology).

3. The event

This is narrated, with great concision, in rr. 50–163:

*wi hafe läger weþ 2 skjar en þags rise norr fro þeno sten. wi war ok fiske en þag.
äptir wi kom hem, fan 10 man röþe af bloþ og þep. AVM fräelse af illy.
“we had camp by 2 skerries (shelters?¹²) one day’s journey northward from
this stone. We went (lit. were) fishing one day. Afterwards we came home,
found ten men red with blood and dead. A[ve] V[irgo] M[aria], Save [us] from
evil!”.*

In this section, syntax and morphology combine to eliminate redundancy and thus to tighten the “texture” (to use the term suggested by Schiller [1980]) of the message. One kind of redundancy is that which results from the requirement, in the earlier stages of all the Indo-European languages, that there must be agree-

ment in grammatical form between the elements of a phrasal construction involving a head and its modifiers, and in a clause between its predicate and its subject (if any).¹³ Examples of the phrasal redundancy involved are Latin *quarum saevarum dearum*, "of what stern goddesses" (gen. pl. fem.) or *illum bonum hominem* "that good man" (acc. sg. masc.).¹⁴ For subject + predicate agreement, cf. Latin *agricola laborat*, "the farmer plows", versus *agricolae laborant* "the farmers plow".

By no means do all the languages of the earth have such requirements as to grammatical agreement. In Hungarian, for instance, the ordinary noun has a singular and a plural, e.g., *kabát* "overcoat", *kabátok* "overcoats". However, in a phrase involving a numeral plus a noun, agreement in number is not customary. One says, for instance, not **hét kabátok* "seven overcoats", but *hét kabát*, literally "seven overcoat". One indication of plurality is sufficient, and redundancy of number-reference is thus avoided.

There are several sub-types of elimination of redundancy in the Kensington inscription involving reduction of grammatical agreement, similar to but not identical with the Hungarian phenomenon which we have just discussed. In each of these sub-types the initial instance of the construction manifests an inflected form, and the following instances have only the non-inflected "basic" form of the verb or adjective involved. Of these sub-types, two are found in section 2 of the inscription and two in the third section.

A. Verbs, in subject-pronoun + verb constructions. In section 2 (rr. 50-163) there are three clauses, each introduced by the subject-pronoun *wi* "we" and having a verb as the center of the predicate, as shown in the table.

Rune-nos.	Subject and Predicate	Gloss	"Classical" Old Swedish
50-55	<i>wi hape</i>	"we had"	<i>hafdum</i>
96-106	<i>wi war ok¹⁵ fiske</i>	"we were (i.e., went) and fished"	<i>varum fiski</i>
118-128	<i>wi kom [...] fan</i>	"we came" [...], found	<i>komum, fannum</i>

Table: subject + predicate constructions

In this sequence of subjects and predicates, the first verb, *hape*, is inflected for tense (root *ha*-plus suffix *-pe*, past tense) but not for number or person, since these are indicated by the subject pronoun *wi*. In the remaining four verbs, which are all in parallel constructions, there was no need of repeating any inflectional

suffixes to indicate tense, person, or number, since these had already been manifested in the first of these smaller constructions. In this way, the reader/speaker's time and energy are spared, and unstressed syllables are eliminated, thus giving greater prominence to the necessarily stressed root-syllables (cf. below, s.v. Phonology). No violence was done to the forms, since these were undoubtedly already circulating in every-day speech.¹⁶

B. Descriptive adjectives, in the sequence *röþe af bloþ og þed* (rr. 135–147) “red with blood and dead”. Here, the plurality of the two adjectives *röþe* and *þed* is adequately indicated by the *-e* ending of the first adjective, *röþe*, and the second adjective, *þed*, which climaxes the entire narrative section, is thus kept in its basic monosyllabic form, carrying an even greater “punch” than it would have in its inflectional form, and hammering home the conclusion of the entire narrative.

4. Further background

The third section of the text adds further information concerning the situation¹⁷ in which the expedition found itself, an intermediate stage of their trip before they arrived at the island on which they were located, and the date (rr. 164–222):

har 10 mans we hawet at se äptir wore skip 14 þagh rise from þeno öh ahr 1362
 “There are 10 men by the sea to look after our ships 14 days’ journey from this island, year 1362”.

In this section there are two additional instances of the elimination of redundancy through the dropping of a final unstressed syllable, and one more use of an impersonal, subjectless verb:

A. In the construction preposition + verb object *at se* (rr. 179–182) “to see, look”, there was no need to use the infinitive form with an added suffix (as in “classical” Old Swedish *sea*), since the meaning of the phrase is completely covered by the bare root *se-*, and, as in the preceding examples (above, section 2 A), the force of the verb is not diminished as it would have been by the unnecessary addition of a following unstressed syllable.

B. The possessive adjective *wore* (rr. 188–191) “our” signals plural meaning through the ending *-e*, thus rendering unnecessary the indication of plurality in the noun *skip*¹⁸ (rr. 192–195) ‘ships’, with the resultant phrase *wore skip* (rr. 188–195) “our ships”.

C. The verb-form *har* (rr. 164–166) can function as either a regularly inflected verb with a subject, in the meaning “have”, or as an impersonal verb with a

direct object but inherently without a subject,¹⁹ meaning “there is, there are”. Many, perhaps most, commentators have assumed that *har* is being used here in the first of these meanings, and have accordingly wished to posit the ellipsis of a hypothetical subject-pronoun **wi* “we”. Here, however, there is no need to assume such an ellipsis, since *har* is clearly being used as an impersonal “there is, there are”. As we have pointed out with regard to the beginning of the text (cf. above, section 1 A), the presence of a relatively unstressed *wi* would have weakened the effect of the onset of the sentence, depriving it of a more effective beginning with the fully stressed *har*.

5. Phonology

It is generally considered that a phonemically significant contrast between “strong” (i.e., stressed) and “weak” (i.e., unstressed) syllables has characterized all the Germanic languages ever since their constitution as a separate group within the Indo-European family.²⁰ We are therefore justified in assuming that the speech of the Kensingtonian explorers had this same characteristic. Heavy stress, in any Germanic language, conveys semantic emphasis, and weak stress signifies a relative lack of importance. We may, therefore, when we read the text aloud or listen to it “in our mind’s ear” as we read it to ourselves,²¹ have an understanding of the effect it must have had in its own time. The reduction of verbal and adjectival forms was not due to a putative forger’s ignorance of “correct” Old Swedish grammar. The predominance of stressed syllables undoubtedly gave the text a strong cohesion and impressive authority, ending the main narrative with a dramatic conclusion on the ominous syllable *þep*.

Two final questions:

- (1) Of all the persons in the Kensington area in the 1800s, how many would have had the knowledge of this and many other features of late mediaeval Scandinavian languages to concoct a text of this kind? Not a one.
- (2) On the other hand, how many other characteristics of the language of that period remain to be identified in this inscription? There may well be a number awaiting our study, once current inertia and ingrained anti-Kensingtonian prejudices are overcome.

Author’s address

Robert Andersen Hall, Jr. (born 4 April 1911, died 2 December 1997)

Notes

1. For an overall picture of the present situation, cf. especially Hall (1994), Nilsestuen (1994), and Nielsen (1994 and Forthcoming a, b). The two first-mentioned cover much of the same material, with the same basic outlook, but from somewhat different points of view; they complement each other to a considerable extent.
2. The most detailed picture of events surrounding the discovery of the stone by Olof Öhman is given by Nilsestuen (1994: Chapter 2). In view of the well-attested facts of the event, no credence should be given to the wilful distortions and downright fabrications of Wahlgren in his 1958 book, repeated unchanged in many of his later publications. Cf. Hall (1994: Chapter 15); Nilsestuen (1994: Chapter 4).
3. Cf. Nielsen (Forthcoming b).
4. On cohesion as an essential feature of discourse in general, including narrative, cf. Halliday & Hasan (1976) and Schiller (1980).
5. I use "prosodic" here in the sense of 'involving further features beyond those of straight "linear" phonemes, e.g., stress, length, and pitch, as emphasized especially by J. R. Firth and his followers (cf. Christie [1980]).
6. We have no way of knowing whether the text, as we have it at present, was the work of one or of several members of the expedition. Since the inscription is clearly a "fair copy", the latter may well have been the case.
7. I follow here the method of stylistic analysis developed by M. Riffaterre (1959, 1960), distinguishing between stylistic background, context, and device.
8. The status of *AVM* as a lexeme in this inscription is open to question, since we do not know what (if any) correspondence may have existed between this sequence of Roman letters and one or another phenomenon of the everyday speech represented in the rest of the text. There may have been some, as is the case of the use of the abbreviation *Tel.* in Japan for 'telephone' in signs, advertisements, and the like, and read off with the Japanese word for 'telephone'. Or, if there was in the Kensingtonians' speech some Germanic equivalent for 'Hail Virgin Mary', we have no way of knowing what its specific linguistic form may have been.
9. In this context, *göter* may refer either to 'men from Västergötland' (as assumed by almost all commentators) or perhaps to 'men from the island of Gotland' (Hall, 1994: Chapters 4, 5).
10. As suggested by Nielsen (Forthcoming a).
11. In order to assume an "ellipsis", we must first demonstrate that there must have been, at a previous stage of the development of the construction in question, a clearly identifiable element. In various instances, this is not the case, and then we must not postulate an "ellipsis" (cf. note 19, below).
12. Some scholars (e.g., Nielsen forthcoming a) prefer to take *skjar* (rr. 65–69) as referring to 'shelters' rather than to 'skerries'. If the presumed "shelters" were of the Scandinavian "bog-hut" variety, attested elsewhere in North America (e.g., at Spirit Pond in Maine; cf. Lenik [1973]), rather than as mere protections against the wind or rain, such an emendation may not be out of place.

From the long-range point of view, we should begin to take into account the findings of such historians as Mallery & Robinson (1979) and G. Thompson (1992) concerning the possibility of there having been, not only individuals or small parties wandering about, but also longer-lasting settlements of Norse and Irish. For the Norse origin of the place-name

- Norumbega*, attested in maps since colonization times, for what is now the northeast region of the United States, cf. Hall (1996).
13. Cf. most recently Lehmann (1995).
 14. Cf. the discussion of this point by Sapir (1921: 121), characterizing such sequences as "relentless rhymes".
 15. I have kept *ok* (rr. 102–103) 'and' in this sequence because it is a necessary connector in a set phrase (cf. Hall [1994: 45]).
 16. As pointed out by Hagen (1950: 340): "[M]uch human speech along the highways and byways of Sweden [and also of Norway, Denmark, and every other speech-community on earth — RAHjr] must have remained unrecorded".
 17. Might the material on the side of the stone have been added as an after-thought at some stage of the elaboration of the text? (This is pure speculation, but not harmful if recognized as such.)
 18. In *skip* (rr. 192–195) 'ship(s)', rune 195 is very clearly the *p*-rune, not the *b*-rune, and should be transcribed with the Roman *p*, not with *b*. It is the function of the middle line in presentations of this type to give, not a phonetic or phonemic representation of what we think it may have been its pronunciation, but a transliteration, with only one Roman equivalent for each rune. Further phonological interpretation belongs in a commentary. (Cf. most recently Barnes [1994] on this point.)
 19. By no means can all "subjectless" sentences be regarded as transforms of sentences which had had subject pronouns but had lost the pronouns through ellipsis ("pro-drop" in currently fashionable terminology). In instances like that of *har* (rr. 164–166) 'there are') on the Kensington stone, or similar verbs standing alone like modern Spanish *hay* 'there is, there are', Old French *a*, etc. of similar meaning, or Latin *pluit* 'it rains', *ninguit* 'it snows', and many others, none can have a subject when used in their literal meaning, not even in the deepest of "deep structures". In all the conservative Indo-European languages, the only "deep structure" that can be regarded as the most fundamental is predicate alone, not subject + predicate, no matter what traditional Latin grammar or Chomskyan orthodoxy may say.
 20. Cf. Haugen (1976: 151): "The continued presence of *initial stress* as a 'forward peak of energy' (Brosnahan & Turner 1958) tended to concentrate attention on the information value of the first syllable", and also (204): "The previously noted trend of G[er]m[ani]c to concentrate information in the stressed syllable and reduce the importance of the rest of the word".
 21. It should be generally known, but is not, that even when we read "silently" we are actually subvocalizing all the time, and repressing the muscular activity of the organs of speech (cf. Edfeldt [1960]).

References

- Barnes, Michael P. 1994. *The Runic Inscriptions of Maeshowe (Orkney)*. Uppsala: Runrön — Runologiska bidrag av Institutionen för nordiska språk vid Uppsala Universitet.
- Brosnahan, L. E. & G. W. Turner. 1958. *The Phonetic Tendency During the Formative Period of the Old Norse Vowel System*. *Arkiv for Nordisk Filologi* 17.119–128.
- Christie, William M., Jr. 1980. *Prolegomena to a Neo-Firthian Linguistics*. Lake Bluff, Ill.: Jupiter Press.

- Edfeldt, Åke. 1960. *Silent Reading and Silent Speaking*. Chicago: University of Chicago Press.
- Hagen, Sivert N. 1950. "The Kensington Runic Inscription". *Speculum* 25:321–356, with plates opposite pp. 321, 332.
- Hall, Robert A., Jr. 1950. "A Syntactic Archaism in the Kensington Rune-Stone". Hall 1982, 99–106.
- . 1982. *The Kensington Rune-Stone is Genuine*. Columbia, South Carolina: Hornbeam Press.
- . 1994. *The Kensington Rune-Stone: Authentic and important*. Lake Bluff, Ill.: Jupiter Press.
- . 1996. "The Etymology of *Norumbega*". *New England Antiquities Research Association Journal* 30:3.91–94.
- Halliday, Michael & Ruqaiya Hasan. 1976. *Cohesion in English*. London: George Allan & Unwin.
- Haugen, Einar. 1976. *The Scandinavian Languages*. Cambridge, Mass.: Harvard University Press.
- Lehmann, W. P. 1995. "Agreement languages". *LACUS Forum* 22:19–58.
- Lenik, Edward J. 1973. *The Spirit Pond Sod House*. Milford N.H.: New England Antiquities Research Association Special Publications no. 2.
- Mallery, Arlington & Mary R. Robinson. 1979. *The Recovery of Lost America: The story of the pre-Columbian Iron Age in America*. New York: Duell, Sloane & Pierce.
- Nielsen, Richard. 1995. "Linguistic Aspects Concerning the Kensington Rune-Stone". Epigraphic Society Occasional Contributions.
- . Forthcoming a. "Runic Aspects Concerning the Kensington Rune-Stone". To appear in Epigraphic Society Occasional Contributions.
- . Forthcoming b. "A Century of Controversy — The Kensington Rune Stone".
- Nilsestuen, Rolf M. 1994. *The Kensington Rune Stone Vindicated*. Lanham, Maryland. & New York: University Press of America.
- Nygaard, M. "Udeladelse av Subjekt: subjektlose Sætninger i det norrøne Sprog (den klassiske Sagastil)". *Arkiv for Nordisk Filologi* 10 (= NF 7.) 1–25.
- Riffaterre, Michael. 1959. "Criteria for Style Analysts". *Word* 12:154–174.
- . 1960. "Stylistic Context". *Word* 13:207–218.
- Sapir, Edward. 1921. *Language*. New York: Harcourt, Brace.
- Schiller, Andrew. 1980. Review of Halliday & Hasan (1976). *Forum Linguisticum* 5:88–94.
- Thompson, Gunnar. 1992. *American Discovery: The real story*. Seattle (Washington): Misty Isles Press.

CHAPTER 39

The Compound Gerund in Early Modern English

Matsuji Tajima
Kyushu University

1. Introductory remarks

As noted in the *OED* (s.v. *-ing*¹ 2), the most notable development in the English gerund has been in its acquisition of certain verbal properties, namely, those of: (1) governing an 'accusative' or direct object (e.g., He practises *writing leading articles*); (2) governing a predicative complement (e.g., the habit of *being late*); (3) being modified by an adverbial adjunct (e.g., the habit of *speaking loosely*); (4) indicating tense and voice by means of compound forms (e.g., after *having written* a letter; the necessity of loving and *being loved*); and (5) taking a common-case or objective-case subject (e.g., I insist upon *Miss Sharp appearing*; what is the use of *me speaking*?). Thus, the English gerund, which began as a pure noun, has broadened its syntactic role beyond anything characteristic of its own history or of that of the other Germanic languages. In my earlier work (Tajima 1985: 111–29), in which I tried to delineate the gerund's acquisition of these verbal features in the Middle English period, I observed in connection with the fourth feature above that the compound passive voice form (e.g., for *being* + past participle) is attested from 1417, but remains only in embryo in the 15th century, whereas the compound perfect form (e.g., for *having* + past participle) does not come into use at all in ME, and developed only in early ModE. Today, it has even developed the perfect passive form (e.g., for *having been* + past participle). Apparently, however, the development of these compound forms has been gradual.

This verbal feature of showing tense and voice by means of compound forms has never been properly explored and described, as far as I know. From the scant evidence collected to date we are only led to surmise that the gerund

has developed this feature towards 1600 (see e.g., Curme 1912: 362; Trnka 1930: 93–94; Jespersen 1948: §209), but we do not really know when and how these compound tense and voice forms have developed in the early Modern English period. This paper attempts to describe their rise and development in early Modern English, especially in 16th C. English. For this purpose I have examined more than 30 authors or works, which will be listed at the end of this paper.

Before going further, I shall give a list of all the instances of the perfect gerund ('*having* + past participle') and of the passive gerund ('*being* + past participle') found in the texts examined, arranged in the order of their composition or publication.

Text (Date of Composition or Publication)	<i>having</i> + pp.	<i>being</i> + pp.	<i>having</i> <i>been</i> + pp.
Skelton, <i>Complete Poems</i> (1499–1522)			
Watson, <i>Valentine and Orson</i> (?1503–5)	2	1	
Fisher, <i>English Works</i> (1509–35)			
Medwall, <i>Fulgens and Lucrece</i> (c1515)			
J. Heywood, <i>Plays</i> (c1525–c1544)			
Parker, <i>Forty Six Lives</i> (1534–47)			
Bale, <i>Dramatic Writings</i> (1537–40)			
Palsgrave, <i>Comedy of Acolastus</i> (1540)			
Ascham, <i>Toxophilus</i> (1545)		2	
Vicary, <i>Anatomie</i> (1548)			
More, <i>Utopia</i> (1551)		1	
Udall, <i>Roister Doister</i> (?c1553)			
Gammer Gurton's <i>Needle</i> (1553)			
Ascham, <i>A Report and Discourse</i> (1553)			
Ascham, <i>The Scholemaster</i> (1570)			
Gascoigne, <i>Supposes</i> (1573)			
Lyly, <i>Euphues</i> (1578)		2	
Sidney, <i>Letter to Elizabeth</i> (1579)		1	
Lyly, <i>Euphues and his England</i> (1580)		3	
Sidney, <i>(Old) Arcadia</i> (1580–81)	7	10	
Sidney, <i>Defence of Poetry</i> (1580)		2	

Text (Date of Composition or Publication)	<i>having</i> + pp.	<i>being</i> + pp.	<i>having</i> <i>been</i> + pp.
Christopher Marlowe (1587–1616)		6	
Nashe, <i>The Anatomie of Absvrditie</i> (1589)			
William Shakespeare (1589–1613)	4	8	
Spenser, <i>Faerie Queene</i> (1590–96)	5	8	
Greene, <i>James the Fourth</i> (1591)			
Kyd, <i>The Spanish Tragedy</i> (1592)			
Nashe, <i>Pierce Penilesse</i> (1592)			
Nashe, <i>The Vnfortunate Traveller</i> (1594)		2	
Deloney, <i>Novels</i> (?1597–98)			
Dekker, <i>The Shoemaker's Holiday</i> (1599)		1	
Ben Jonson (1601–14)		9	
T. Heywood, <i>Kindness</i> (1603)			
Total	18	56	0

(The figure 0 (zero) has been omitted in the above table.)

2. The compound perfect gerund

The gerund had originally, and to a great extent still has, no time reference. Accordingly, the simple, non-compound *-ing* form was employed to express any time or no time in particular, as in: "He is very fond of *reading* novels" (referring to the present time); "I remember *seeing* him once" (the past); "I intend *seeing* him" (the future); "*Seeing* is *believing*" (no time). Later in its development, the gerund began to appear with compound forms (consisting of *having* plus the past participle of the verb) for the perfect tense, denoting the past time-sphere.

Previous studies have suggested that the perfect gerund comes into existence as late as early ModE, more specifically towards the close of the 16th C. (see e.g., Curme 1912: 362; Trnka 1930: 93–94; Mustanoja 1960: 573). As stated before, it is non-existent in the ME period (Tajima 1985: 112). Even in the early ModE period, it is never used by More (Visser 1946: 442), and only exceptionally by Spenser (Sugden 1936: 143–44) and Shakespeare (Jespersen 1931: §7.8(2)).

Even as late as the second half of the 17th C. it still seems to have been very rare; John Dryden's prose furnishes only six instances (Söderlind 1958: 20). Visser (1973: 2230–31) attempts to give all the examples collected by earlier scholars for the period up to 1614, beginning with c1528; however, the collection, for the most part, actually consists of those of the perfect participle,¹ only seven instances being relevant ones. As a matter of fact, however, all those previous studies which Visser consulted have revealed 12 relevant instances in all. The first recorded instances do not seem to occur until 1580–81 in Sidney's *Arcadia*, as indicated by Blume (1880: 41). Thus as late as about 1600 the construction with *having* is still very rare, the simple form remaining the rule, as in: Shakespeare, *Two Gentlemen* IV.i.45–46 "Myself was from Verona banished For practicing to steal away a lady".

Altogether, I have found 18 relevant examples in my corpus, out of which seven are the same as those already noted by Blume (1880: 40–41), OED (s.v. *-ing*¹ 2), Trnka (1930: 94), Jespersen (1931: §7.8(2)), Sugden (1936: 143–44), Biese (1950: 7–8) and Visser (1973: 2230–31). As the table above clearly shows, it is not used in most of the 16th C. texts I have examined, appearing only twice in Watson, seven times in Sidney, five in Spenser, and four in Shakespeare. The earliest examples that I have found are in Watson's *Valentine* and *Orson* dated about 1503–5, although no other examples seem to occur in the intervening 80 years between Watson and Sidney. In view of the importance of the earliest stage of development of this construction, all of the examples collected for the period up to 1614 are quoted below:

- (1) ?1503–5 Watson, *Valentine* 98/3–6 I muste doo a message to kyng
Pepyn of Fraunce from a sister of his named Bellyssant that of long
time was banyshed out of Constantynoble, with wronge, and without
hauynge deserued it.
- (2) ?1503–5 Watson, *Valentine* 147/14–16 I wolde not lyue to se suche
two valyaunt knyghtes dye before me without *hauynge deserued* it.
- (3) 1580–81 Sidney, (*Old*) *Arcadia* 20/28–30 I doubt not I shall quickly
have occasion rather to praise you for *having conquered* it than to
give you any further counsel how to do it.
- (4) 1580–81 Sidney, (*Old*) *Arcadia* 65/5–8 after *having railed* at the
name of Cupid as spitefully as he could devise, ... he thus, tuning
his voice to a rebeck, sang against him.
- (5) 1580–81 Sidney, (*Old*) *Arcadia* 96/10–13 he ... thanking his wife for
having entertained Cleophila, desired her she would now return into
the lodge.

- (6) 1580–81 Sidney, *(Old) Arcadia* 224/32–33 ... whose pillow she kissed a thousand times for *having borne* the print of that beloved head.
- (7) 1580–81 Sidney, *(Old) Arcadia* 279/15–17 Her reason began ... to tear itself with anguish for *having made* so weak a resistance.
- (8) 1580–81 Sidney, *(Old) Arcadia* 301/24–25 But if it be to bring me to any punishment whatsoever for *having undertaken* so unexcusable presumption.
- (9) 1580–81, Sidney, *(Old) Arcadia* 379/30–33 To me ... what can be more convenient than at least to be at peace with myself in *having discharged* my conscience in a most behoveful verity.
- (10) 1590 Spenser, *Faerie Queene* I.6.42 Yet ill thou blamest me, for *hauing blent* My name with guile and traiterous intent.
- (11) 1590 Spenser, *Faerie Queene* III.2.50 And after *hauing whispered* a space Certaine sad words, with hollow voice and bace, She to the virgin said.
- (12) 1590 Spenser, *Faerie Queene* III.5.33 And after *hauing searcht* the intuse deepe, She with her scarfe did bind the wound from cold to keepe.
- (13) 1590 Spenser, *Faerie Queene* III.7.29 And after *hauing* him *embowelled*, ... it chaunst a knight To passe that way, as forth he trauelled.
- (14) 1596 Spenser, *Faerie Queene* VI.2.39. And after *hauing* them vpon him *dight*, He tooke that Ladie, and her vp did rayse.
- (15) 1592–93 Shakespeare, *Venus and Adonis* 809–10 Mine ears ... Do burn themselves for *having* so *offended*.
- (16) 1594 Shakespeare, *The Two Gentlemen* I.iii.15–16 Which would be great impeachment to his age, In *having known* no travel in his youth.
- (17) 1607–8 Shakespeare, *Coriolanus* V.iii.113–18 thou Must ... bear the palm for *having* bravely *shed* Thy wife and children's blood.
- (18) 1611 Shakespeare, *Tempest* III.i.18–19 When this burns, 'Twill weep for *having wearied* you.

Out of these 18 examples, only seven, namely (10)–(12) and (15)–(18), have been mentioned previously. Below are five examples noted elsewhere but which were not found in my corpus:

- (19) 1580–81 Sidney, (*New*) *Arcadia* (*Prose Wks.*, ed. Feuillerat) 37/3–4 He should have suffered death for *having slaine* the Kings Nephew. (Quoted from Blume 1880: 41)
- (20) 1580–81 Sidney, (*New*) *Arcadia* 51/15 after *having failed* to take him into the fisher boate, he had ... come to this Gentlemans house. (Visser 1973: 2230)
- (21) 1580–81 Sidney, (*New*) *Arcadia* 66/3–4 I condemne my selfe of want of consideration in not *having demaunded* thus much. (OED s.v. *-ing*¹ 2 = Trnka 1930: 94 = Visser 1973: 2230)
- (22) 1580–81 Sidney, (*New*) *Arcadia* 150/5–6 At the first word (thanking his wife for *having entertained* Zelmane,) he desired her she would now returne into the lodge. (Blume 1880: 41)
- (23) 1595 Spenser, *Colin Clout Comes Home Again*e 904 That ill ... they him requite, For *having loved* ever one most dear. (Visser 1973: 2231)

Judging from my own findings and the evidence found elsewhere, the compound perfect gerund occurs much earlier than has previously been noted, appearing from about the beginning of the 16th C. — the earliest examples dating from ?1503–5 Watson, *Valentine and Orson*, as given in (1) and (2) above — but it is still very rare in the 16th C. and even until the early 17th C. It began to appear sporadically in the last quarter of the 16th C. but exclusively in Sidney, Spenser, and Shakespeare, as far as my corpus is concerned.

3. The compound passive gerund

The gerund was also originally, and to a great extent still remains, neutral as regards voice, that is, as far as the distinction between the active and the passive voice is concerned. In consequence, the simple or active *-ing* form also had to serve a passive function until a new passive form developed, as in Shakespeare, *Merry Wives of Windsor* II.v.9 “I could do nothing without *bidding*”; *Hamlet* II.ii.532 “and who should scape *whipping*”. This inherent, nominal indifference to voice has survived in present-day English in more or less set phrases after *need*, *want*, *deserve*, *worth*, *past*, and a few others (see Jespersen 1940: §§9.2.3–9.2.4), as in “My shoes need *mending*”; “The book is worth *reading*”. In the course of time, however, a new passive compound form (consisting of *being* plus the past participle of the verb) has come into use for passive import. Although it is commonly held that this new passive gerund makes its first

appearance about 1600, Mossé (1938: 148) records three earlier instances from the 15th C., the earliest dated 1417:

1417 Ellis *Orig. Lett.* II i 59 but now your sayd leiges, both their and elsewhere, may suffer their goods and cattels to remayne in the feilds day and night without *being stolen*; c1454 Pecock. *Fol.* 126/26–29 þe instrument is not wircher of þe same actyue deede principali and þou3 his owen strengþ in *being restid*, but in strengþ of an opir, and in *being movid*.

To these ME examples I have added two more instances (Tajima 1985:114):

c1454 Pecock *Fol.* 126/14–17 þou3 þe stoon have a wirching, 3it he wirchip not and doop not þilk wirchyng bi him silf and bi his owen strengþ, as in a *being restid*, but in strengþ of þilk vertu, and in a *beyng movid*.²

These five are all the examples so far found for the ME period.

Although it appeared much earlier in writing than the compound perfect gerund, the compound passive gerund apparently remained scarce until about 1600, as is shown in the table in §1. It is non-existent in the great majority of texts written or published before 1550, and even in many of the texts after 1550, but it does occur occasionally in the last quarter of the 16th C., namely in Lyly, Sidney, Spenser, Shakespeare, Jonson, and Marlowe. In order to demonstrate its slow progress in early ModE, all the relevant examples I have collected are quoted or mentioned below:

- (24) ?1503–5 Watson, *Valentine* 315/14–16 In this wyse was Valentyne a great whyle within his palays without *beyng known* of any body.
- (25) 1545 Ascham, *Toxophilus* 52/19–21 whiche shoulde take more honor in *being coupled* to Englande, then we shulde take profite in *being ioyned* to Scotlande.
- (26) 1551 More, *Utopia* 120/37–121/1 Yea, they saue it as muche as maye be from *beinge ouerrune* and *troden* downe, other with men or horses.
- (27) 1578 Lyly, *Euphues* 238/18–19 whosoeuer accompteth you a lyar for praising me, may also deeme you a letcher for *being enamoured* of me.
- (28) c1579 Sidney, *A Letter to Queen Elizabeth* 47/15–16 ... nor inward weakness doth lightly subvert itself without *being thrust* at by some outward force.
- (29) 1580 Lyly, *Euphues and his England* 27/13–14 *Callimachus* obstinate in his fond conceit, was so far from *being perswaded* by this old Hermit. (Also in: 6/15; 139/24)

- (30) 1580 Sidney, *A Defence of Poetry* 115/22–23 we are ravished with delight to see a fair woman, and yet are far from *being moved* to laughter. (Also in: 91/15)
- (31) 1580–81 Sidney, *(Old) Arcadia* 74/4–6 O noble age who place their only bliss In *being heard* until the hearer die, Utt'ring a serpent's mind with serpent's hiss! (Also in: 112/9; 115/4; 157/29; 203/24; 209/3; 288/9; 308/24–25; 376/32; 405/28)
- (32) 1587 Marlowe, *Tamburlaine, Part II* IV.i.130–31 Wherat thou trembling hiddest thee in the air, Cloth'd with a pitchy cloud for *being seen*. (Also in: c1587 *Jew of Malta* II.iii.108; 1594 *Dido* V.i.145; 1598 *Hero and Leander* II.36; 1616 *Dr. Faustus* iii.82, iii.86)
- (33) 1589–90 Shakespeare, *1 Henry IV* III.ii.46–47 By *being seldom seen*, I could not stir But like a comet I was wond' red at. (Also in: 1602–3 *All's Well That Ends Well* I.I.130; 1604 *Othello* I.iii.137; 1607–8 *Pericles* I.ii.22; 1609–10 *Cymbeline* II.iii.131; 1611 *The Tempest* I.ii.91, II.i.147)
- (34) 1590 Spenser, *Faerie Queene* III.4.50 Yet former feare of *being fowly shent*, Carried her forward with her first intent. (Also in: III.4.58; 1596 *Faerie Queene* V.2.16, V.5.44, V. 7.21, VI.6.18, VI.8.26, VI.10.35)
- (35) 1594 Nashe, *The Vnfortunate Traveller* 303/4–5 The diuel and I am desperate, he of *being restored* to heauen, I of *being recalled* home.
- (36) 1599 Dekker, *The Shoemaker's Holiday* V. ii. 51 O pardon me, dear love, for *being misled*.
- (37) 1603 Ben Jonson, *Sejanus* I.499–501 only a long, A lasting, high, and happy memory They should, without *being satisfied*, pursue. (Also in: 1603 *Sejanus* IV.9, IV.23; 1605 *Volpone* III.vii.41;³ 1610 *The Alchemist* III.iv.124; 1614 *Bartholomew Fair* IV.vi.15–6, V.ii.67)

Thus it is apparent that the compound passive gerund, though definitely occurring more often and more widely than the compound perfect gerund, remained scarce in the 16th C. and even as late as the early 17th C.

4. The compound perfect passive gerund

Apparently the perfect passive gerund (consisting of *having been* plus the past participle of the verb) developed still later. It is not attested in my early Modern

English corpus. This particular compound form seems to have received hardly any attention in the past. Visser (1973: 2431–32) makes no mention of it, although he says that the perfect passive *participle* first appears in the last quarter of the 16th C. (from 1590 Sidney, *Arcadia*). Kisbye (1992: 133) suggests that it begins to occur only in the 18th C., but gives no factual evidence.⁴ Only Jespersen (1931: §7.8(4)) quotes four such examples, beginning with 1837 Dickens *Pickwick Papers*. Incidentally I have noted some slightly earlier examples in Jane Austen's novels: three in *Sense and Sensibility* (1811), two in *Pride and Prejudice* (1813), one in *Emma* (1816), two in *Northanger Abbey* (1818), and three in *Persuasion* (1818). For reference purposes, some examples are quoted below:

- (38) 1811 Austen, *Sense and Sensibility* 188/21–23 But her condemnation of him did not blind her to the impropriety of their *having been written* at all.
- (39) 1813 Austen, *Pride and Prejudice* 338/14–16 Bingley looked a little silly at this reflection, and said something of his concern, at *having been prevented* by business.
- (40) 1837 Dickens, *Pickwick Papers* 54 certificates of her *having been brought up* in the way she should go when young. (Jespersen)

How far back the construction in question can be traced in writing remains to be seen, but it appears to be current in the 19th C., as the above examples suggest.

5. Concluding remarks

This paper has been concerned with the assumption of compound tense and voice forms by the gerund in early ModE. It has indicated that the compound perfect gerund, which was not attested in ME, occurs as early as about 1503–5, some 80 years before the oldest instance on record (1580–81 Sidney *Arcadia*), but that it remains extremely rare throughout the 16th C. and even beyond. On the other hand, the compound passive gerund, which was evidenced as early as 1417 (Ellis *Orig. Lett.*) but remained in embryo throughout the 15th C., appears more often than the perfect gerund, but it also makes slow progress even as late as 1600. Moreover, the perfect passive gerund remains totally undeveloped in early ModE. From this survey it can be concluded that the compound gerund (perfect or passive) was far from being well developed before 1600 or even in the early 17th C. When it became a regular feature of the English language still remains to be explored.

Acknowledgments

I am grateful to Dr. Peter Rawlings and Ms. Nobuko Suematsu for their helpful comments on an earlier version of this paper.

Author's address

Matsuji Tajima
Department of English
Kyushu University
Ropponmatsu, Fukuoka 810-8560, JAPAN
tajima@rc.kyushu-u.ac.jp

Notes

1. According to Biese (1950: 3-18) and Visser (1972: 2230-31), the compound perfect participle, though occurring first in the early 16th C., appears to have been a fairly well established idiom by Shakespeare's time.
2. It is to be noted, however, that, although the two instances from *Pecock* display strong verbal force (because of the compound voice form employed), the fact that each is preceded by an article makes it impossible to ignore the nominal content involved.
3. Note that the old, active and the new, passive form occur in juxtaposition in: 1605 Jonson, *Volpone* III vii.40-41 "What is my gold The worse for *touching*? clothes for *being looked on*?"
4. See also Kisbye (1971: 65), in which he notes that it is not current till the 19th C.

References

Part I: Primary Sources

- Ascham, Roger. 1904. *English Works of Roger Ascham* ed. by W. A. Wright. Cambridge: Cambridge University Press (Repr., 1970). ("Toxophilus" [1545]; "A Report of the Affaires and State of Germany" [1553]; "The Scholemaster" [1570]).
- Bale, John. 1907. *The Dramatic Writings of John Bale* (1537-40) ed. by John H. Farmer. London: Early English Drama Society.
- Deloney, Thomas. 1961. *The Novels of Thomas Deloney* (?1597-98) ed. by Merritt E. Lawlis. Bloomington. Ind.: Indiana University Press.
- Dekker, Thomas. 1927. *The Shoemaker's Holiday* (1599) ed. by Wilfrid J. Halliday. London: Macmillan.
- Fisher, John. 1876. *The English Works of John Fisher* (1509-35). EETS ES 27.

- Gammer Gurton's Needle* (1553). 1970 (1934¹) *Five Pre-Shakespearean Comedies* ed. by Frederick S. Boas. London: Oxford University Press.
- Gascoigne, George. 1970 (1934¹). *Supposes* (1573). *Five Pre-Shakespearean Comedies* ed. by Frederick S. Boas. London: Oxford University Press.
- Greene, Robert. 1970. *The Scottish History of James the Fourth* (c1591) ed. by Norman Sanders (The Revels Plays). London: Methuen.
- Heywood, John. 1991. *The Plays of John Heywood* (c1525–c1544) ed. by Richard Axton & Peter Happé. Cambridge: D.S. Brewer.
- Heywood, Thomas. 1961. *A Woman Killed with Kindness* (1603) ed. by R. W. Van Fossen (The Revels Plays). London: Methuen.
- Jonson, Ben. 1969. *Everyman in His Humour* (1601) ed. by G. B. Jackson (The Yale Ben Jonson). New Haven. Conn.: Yale University Press.
- . 1983. *Volpone, or, The Fox* (1607) ed. by R. B. Parker (The Revels Plays). Manchester: Manchester University Press.
- . 1990. *Sejanus His Fall* (1607) ed. by Philip Ayres (The Revels Plays). Manchester: Manchester University Press.
- . 1967. *The Alchemist* (1610) ed. by F. H. Mares (The Revels Plays). London: Methuen.
- . 1960. *Bartholomew Fair* (1614) ed. by E. A. Horsman (The Revels Plays). London: Methuen.
- Kyd, Thomas. 1959. *The Spanish Tragedy* (1592) ed. by Philip Edwards. (The Revels Plays). London: Methuen.
- Lyly, John. 1902. *The Complete Works of John Lyly*, Vols. I & II, ed. by R. Warwick Bond. Oxford: Clarendon Press. (Vol. I: *Euphues* [1578]. Vol. II: *Euphues and his England* [1580])
- Marlowe, Christopher. 1978. *The Jew of Malta* (c.1587) ed. by N. W. Bawcutt (The Revels Plays). Manchester: Manchester University Press.
- . 1981. *Tamburlaine The Great* (1587) ed. by J. S. Cunningham (The Revels Plays). Manchester: Manchester University Press.
- . 1968. 'Hero and Leander'. in *The Poems* (1598) ed. by Millar MacLure (The Revels Plays). London: Methuen.
- . 1968. *Dido Queen of Carthage* (1594) and *The Massacre at Paris* (c.1589) ed. by H. J. Oliver (The Revels Plays). London: Methuen.
- . 1962. *The Tragical History of the Life and Death of Doctor Faustus* (1616) ed. by John D. Jump (The Revels Plays). London: Methuen.
- Medwall, Henry. 1970 (1934¹). *Fulgens and Lucrece* (c1515). in *Five Pre-Shakespearean Comedies* ed. by Frederick S. Boas. London: Oxford University Press.
- More, Sir Thomas. 1904. *Utopia* (Ralph Robynson's transl. 1551) ed. by J. C. Collins. Oxford: Clarendon Press.
- Nashe, Thomas. 1966. *The Works of Thomas Nashe*. Vols. I & II ed. by Ronald B. McKerrow. Oxford: Basil Blackwell. (Vol. I: "The Anatomie of Absvrditie" [1589]; "Pierce Penilesse" [1592]. Vol. II: "The Vnfortunate Traveller" [1594].)

- Palsgrave, John (transl.). 1937. *The Comedy of Acolastus* (1540) ed. by P.L. Carver. EETS 202.
- Parker, Henry (Lord Morley) (transl.). 1943. *Forty-Six Lives from Boccaccio* (1534–47) ed. by Herbert G. Wright. EETS 214.
- Shakespeare, William. 1974. *The Riverside Shakespeare* (1589–1613) ed. by G. Blake-more Evans. Boston: Houghton Mifflin Company.
- Sidney, Sir Philip. 1973. *Miscellaneous Prose of Sir Philip Sidney* ed. by Katherine Duncan-Jones & Jan Van Dorsten. Oxford: Clarendon Press. ("A Letter to Queen Elizabeth" [1579]; "A Defence of Poetry" [1580])
- . 1973. *The Countess of Pembroke's Arcadia (The Old Arcadia)* (1580–81) ed. by Jean Robertson. Oxford: Clarendon Press.
- Skelton, John. 1983. *John Skelton: The Complete English Poems* (1479–1522) ed. by John Scattergood. New Haven. Conn.: Yale University Press.
- Spenser, Edmund. 1909. *Spenser's Faerie Queene*. 2 vols. ed. by J.C. Smith. Oxford: Clarendon Press.
- Udall, Nicholas. 1970 (1934¹). *Ralph Roister Doister* (?c1553 [pub. 1566]). in *Five Pre-Shakespearean Comedies* ed. by Frederick S. Boas. London: Oxford University Press.
- Vicary, Thomas (transl.). 1888. *The Anatomie of the Bodie of Man* (1548) ed. by Frederick J. Furnivall & Percy Furnivall. EETS ES 53.
- Watson, Henry (transl.). 1937. *Valentine and Orson* (?1503–5) ed. by Arthur Dickson. EETS 204.

Part II: Secondary sources

- Biese, Yrjö Moses. 1950. "Notes on the Compound Participle in the Works of Shakespeare and his Contemporaries". *Annales Academiae Scientiarum Fennicae* B.63:2.3–18.
- Blume, Rudolf. 1880. *Über den Ursprung und die Entwicklung des Gerundiums im Englischen*. (Jena diss.) Bremen: Heinrich Frese.
- Curme, G. O. 1912. "History of the English Gerund". *Englische Studien* 45.349–80.
- Jespersen, Otto. 1931–40. *A Modern English Grammar on Historical Principles*. IV–V. Copenhagen: Ejnar Munksgaard; rpt. London: George Allen and Unwin, 1954.
- . 1948. *Growth and Structure of the English Language*. 9th ed. Oxford: Basil Blackwell. (First ed., 1905.)
- Kisbye, Torben. 1971. *An Historical Outline of English Syntax*. Part I. Aarhus: Akademisk Boghandel.
- . 1992. *A Short History of the English Language*. Aarhus: Aarhus University Press.
- Mossé, Fernand. 1938. *Histoire de la forme périphrastique 'être + participe présent' en germanique. II: Moyen-anglais et anglais moderne*. Paris: C. Klincksieck.

- Mustanoja, Tauno Frans. 1960. *A Middle English Syntax*. Part I: *Parts of Speech*. Helsinki: Société Néophilologique.
- OED = *The Oxford English Dictionary* ed. by James A. H. Murray et al. Oxford: Clarendon Press, 1933.
- Söderlind, Johannes. 1958. *Verb Syntax in John Dryden's Prose*. Part II. Uppsala: Lundequistska Bokhandeln.
- Sugden, Herbert W. 1936. *The Grammar of Spenser's 'Faerie Queene'*. Philadelphia: Linguistic Society of America.
- Tajima, Matsuji. 1985. *The Syntactic Development of the Gerund in Middle English*. Tokyo: Nan'un-do.
- Trnka, Bohumil. 1930. *On the Syntax of the English Verb from Caxton to Dryden*. Prague: Jednora Československých Matematiku a Fysiku.
- Visser, Frederikus Theodorus. 1946. *A Syntax of the English Language of St. Thomas More*. Vol. I. Louvain: Uystpruyst.
- . 1973. *An Historical Syntax of the English Language*. Part III, Second Half. Leiden: E. J. Brill.

CHAPTER 40

Gender Affixes in Tsez

Synchrony and diachrony

Bernard Comrie

Max Planck Institute for Evolutionary Anthropology

One of the touchstones in the history of linguistics has been the relationship between synchronic and diachronic investigations. There have been times when it has been argued that the only true explanations are historical. Conversely, there have been periods when synchrony has ruled, with diachrony being viewed at best as an epiphenomenon of the relation between different diachronic states. In this paper, I have chosen one very limited set of data — markers of male and female gender in Tsez — to provide an illustration of how intricate the relationship between synchrony and diachrony can be.

Tsez, also known by the Georgian name Dido, is a member of the Tsezic subgroup of the Nakh-Daghestanian (Northeast Caucasian) language family. Other Tsezic languages are Hunzib, Hinukh, Bezhta, and Khvarshi, which will also figure in passing in the discussion below. Tsez itself is divided into about four main dialect groups: Kidero (*kidiro*), Asax (*asaq*), Mokok (*newo*), and the aberrant Sagada (*sahada*); the first three of these at least are further divided into subdialects, e.g., the Cebari (*ceboru*) subdialect of Asax. The traditional lingua franca of the Tsezic-speaking area in Daghestan is Avar, a member of the Avar-Andi subgroup of the Nakh-Daghestanian language family. Tsezic and Avar-Andi are usually viewed as constituents of an Avar-Andi-Tsez group, although details of subgrouping within Nakh-Daghestanian may change as our knowledge of the languages improves.

Most Nakh-Daghestanian languages have a system of noun classes. In Tsez, there are four classes, conventionally represented by roman numerals. Class I contains all and only words denoting male humans (and human-like entities, such as male supernatural beings). Class II contains all words denoting female humans, though also a fair number of words denoting inanimate objects. Class III

contains all words denoting animals, plus a large number of other words. Class IV contains only inanimate nouns. While the number of noun classes differs from one Nakh-Daghestanian language to another, the distinction between a class centered on male humans and a class centered on female humans is almost universal. It is this distinction that I am here calling "gender".

In Tsez, there is in general no mark on the noun itself that indicates its noun class. Noun class is marked primarily by means of agreement prefixes that appear on most vowel-initial verbs (showing agreement with the intransitive subject or transitive direct object) and adjectives, plus some adverbs and particles. In addition, some pronouns and numerals show agreement by means of more complex morphophonemic alternations. Verb and adjective agreement for class I (prefix \emptyset -) and class II (prefix y -) is illustrated in (1)–(4); this much of the system is common to all Tsezic languages. In Tsez, the class agreement prefix for class III is b -, for class IV r -.

- (1) *uži* \emptyset -*ik'i-s*.
boy I-go-PASTWITNESSED
"The boy went."
- (2) *kid* y -*ik'i-s*.
girl II-go-PASTWITNESSED
"The girl went."
- (3) \emptyset -*exora uži*
I-long boy
"tall boy"
- (4) y -*exora kid*
II-long girl
"tall girl"

The question of whether there is any (synchronic) relation between the form of a noun and its class membership in Tsez is controversial. Rajabov (in press) argues that, while there may historically have been such a relationship, with the initial segment of the noun correlating with the corresponding class prefix, there is synchronically no valid correlation, class membership being determined rather by semantic associations. Comrie & Polinsky (1999) and Polinsky & Jackson (1999) argue that at least certain initial and final segments do correlate synchronically with class membership, in particular that a final or initial high vowel (i , u) or glide (y , w) correlates with class II, in particular for inanimate nouns. It will be recalled that for animate nouns in Tsez there is no controversy concerning the semantic transparency of class assignment: males to class I, females to class II.

This situation differs somewhat from that in Avar, the traditional lingua franca of Tsezic-speakers. Avar has a three-class system with purely semantic assignment: class I (agreement affix *w*) is for male humans, class II (agreement affix *y*) is for females, and class III (agreement affix *b*) is for nonhumans. In Avar, depending on the class of word, agreement affixes can be either prefixes (e.g., on vowel-initial verbs) or suffixes (e.g., on adjectives). However, in Avar there are a number of gender pairs related transparently by the use of the same affixes as mark agreement, i.e., *w* for male human, *y* for female human, as in (5):

- (5) *wás* "boy"; *yás* "girl"
wác'al "cousin (male)"; *yác'al* "cousin (female)"
baḥáraw "bridegroom"; *baḥáray* "bride"

The last item in (5) is in origin adjectival, cf. *baḥárab* "young"; it is conventional to cite Avar adjectives in the class III form. Agreement of an Avar adjective is illustrated in (6)–(7):

- (6) *bércina-w wás*
 beautiful-I boy
 "handsome boy"
- (7) *bércina-y yás*
 beautiful-II girl
 "beautiful girl"

In the body of this paper, I wish to examine some of the finer details of the expression of gender in Tsez, with some comparisons with other Tsezic languages, and concentrating on the interaction between synchronic and diachronic factors. Particular questions that will be the focus of interest are the following: Apart from agreement prefixes, is there a correlation between female human semantics and the phoneme *y*? Is there a correlation between male human semantics and any segment? It will be recalled that the class I agreement prefix in Tsez is zero, though comparison with Avar, where the corresponding segment is *w*, might suggest that we examine *w* as well. I will suggest below that there are two sets of diachronic factors that need to be taken into account: first, the earlier history of the Tsezic languages, which may have left some residues in particular languages; second, the adstratal influence of Avar, whose at times overt use of class affixes on nouns may have influenced the Tsezic languages.

Tsez has borrowed a number of nouns from Avar, including some pairs which bear the Avar class affixes *w* versus *y*, as illustrated in (8) (cf. (5)):

- (8) *wac'al* "male cousin"; *yac'al* "female cousin"
baharaw "bridegroom"; *baharay* "bride"

Although this relation is not, to the best of my knowledge, productive for nouns in any variety of Tsez or any other Tsezic language, it is clear that speakers of Tsez and other Tsezic languages have lexical sets that would enable them to identify an opposition between male *w* and female *y* in nouns. Female *y* will of course be reinforced by the already clearly abstractable class prefix *y* in Tsezic languages; male *w*, however, does not obviously correspond to any abstractable affix in the native vocabulary, since the Tsez(ic) correspondent is \emptyset .

Tsez and other Tsezic languages have also borrowed a number of Avar adjectives, including adjectives bearing class suffixes. In Tsez and most Tsezic languages, such adjectives are, however, invariable. In Tsez, for instance, corresponding to Avar examples (6)–(7), we have *bercinaw uži* "handsome boy", *bercinaw kid* "beautiful girl". There is, incidentally, dialectal variation as to whether the invariable form is the Avar class I (giving *bercinaw*) or class III (giving *bercinab*), but in no dialect of Tsez are these forms described as differentiated for noun class. Indeed, according to Imnajšvili (1963: 47), Khvarshi is the only Tsezic language in which adjectives borrowed from Avar show the Avar class distinction, using the same morphemes as in Avar, e.g., Khvarshi *bercina-w uža* "handsome boy" but *bercina-y kid* "beautiful girl". Khvarshi is geographically separated from the other Tsezic languages by a high mountain range, and is in much more intensive contact with Avar and the Andic languages (themselves under strong Avar influence) than are the other Tsezic languages. The productive use of the Avar agreement suffixes in adjectives, including crucially class I in *-w*, in Khvarshi is thus to be attributed primarily to Avar influence.

There is one other set of adjectives, this time indigenous, where Tsez shows variation between final *w* and *y*, and indeed also *yw*, namely derived adjectives with the suffixes *-šay* (*-šaw*, *-šayw*) "containing, possessing" and *-tay* (*-taw*, *-tayw*) "lacking", e.g., *ciyo-šay* "containing salt", *ciyo-tay* "lacking salt". According to Imnajšvili (1963: 70), the variants are dialectally conditioned, and do not reflect the class of the noun they modify. However, Abdullaev (1981: 91) says that in the Asax dialect, the usual variant of the latter is *-taw*, while *-tay* is used in relation to females.¹ (Note that the forms in *-w* are used everywhere else, rather than being specifically male.) If one assumes that the class distinction is original, then Asax is conservative and other dialects have lost the class distinction here, selecting one or other variant (or the combination in *yw*) as an invariable form. If one assumes that the variation does not reflect an original class distinction, perhaps with an original **yw* being simplified to *y* or *w* in

individual dialects, then the Asax differentiation is an innovation, reflecting the restriction of a variant in terms of its female-correlating final *y*.

But perhaps the most interesting evidence comes from some of the kin terms in the Tsezic languages. In Table 1, selected kin terms in various Tsez dialects are presented, together with corresponding items in other Tsezic languages. A line for a particular entry indicates simply that the language in question does not have a cognate of the Tsez entry. Thus, Tsez is alone among the Tsezic languages in having reflexes of Proto-Nakh-Daghestanian **ʔānV(jV)* “mother”, although the item is attested in other Nakh-Daghestanian languages and is thus presumably a survival rather than an innovation in Tsez (Nikolayev & Starostin 1994: 201–202). Bezhta lacks cognates of Tsez *xediyu*, etc. “husband”; the traditional Bezhta term was *nuzo*, but this is now largely replaced by *bet’eryan*, a loan from Avar *bet’ëryan* “owner, head of household” (Xalilov 1995) — the word *bet’eryan* is used in Tsez in these other senses. Forms are given for different Tsez dialects and from different sources. Some of these differences reflect genuine dialectal or subdialectal correspondences, but others more probably reflect variations even within a single idiolect. For instance, I have texts where the same speaker varies between *eni* and *eniw* “mother”. In some cases other factors condition variation. Thus, in the Cebari subdialect the word for “father” in isolation is usually *obi*, but the dvandva compound for “parents” can appear as either *eni-obi* or *eni-obu*, although *obu* is not possible in isolation; in this same dialect, the word for “mother” in isolation is *eniy*, but in the dvandva compound it appears as *eni-*. The list in Table 1 also indicates the oblique stem where not identical to the citation form, and the plural where not formed from the oblique stem.

Table 1: Selected kin terms in Tsez and other Tsezic languages

	father	mother	husband	sibling
Kidero (Bokarev 1959)				
	<i>obi</i> <i>yu</i>	<i>eni</i> <i>yu</i>	<i>xedi</i> <i>yu</i>	<i>es(i)</i> <i>yu</i> (PL <i>esnabi</i>)
Kidero (Imnajšvili 1963)				
	<i>obi(w)</i>	<i>eniw</i> , <i>enu</i>	<i>xediw</i> (<i>xedyo-</i>)	<i>essu</i> (PL <i>es(na)bi</i>)
Asax (Cebari subdialect; Rajabov MS)				
	<i>obi</i> <i>y</i>	<i>eni</i> <i>y</i>	<i>xedi</i> <i>y</i> (<i>xedyo-</i>)	<i>esi</i> <i>y</i> (PL <i>esnabi</i>)
Mokok (Xalilov MS)				
	<i>obu</i>	<i>eni</i> (<i>enyo-</i>)	<i>xedi</i> <i>yu</i>	<i>esy</i> <i>u</i> (PL also <i>esnabi</i>)
	<i>bab(i)</i> <i>yu</i>	<i>en</i> <i>yu</i>		
Mokok (selected other forms from Abdullaev texts)				
		<i>eniw</i>	<i>xediw</i>	<i>esiw</i> “brother” <i>esiy</i> “sister”

Bezhta (Xalilov 1995)

<i>abo</i>	<i>iyó</i>	—	<i>is</i> (ist'i-) "brother"
			<i>isi</i> "sister"

Hunzib (van den Berg 1995)

<i>abu</i>	—	—	<i>is</i>
------------	---	---	-----------

Hinukh (Nikolayev & Starostin 1994)

<i>obu</i>	—	<i>xeddo</i>	<i>essu</i>
------------	---	--------------	-------------

Khvarshi (Nikolayev & Starostin 1994)

<i>obu</i>	—	<i>xindu</i>	<i>es</i>
------------	---	--------------	-----------

One feature that is immediately striking in these kin terms is the variation, typically with no concomitant semantic differentiation, in the final segments. For instance, the Tsez word for "father" appears as *obi*, *obiy*, *obiw*, and *obiyu* (plus *-obu* in the dvandva compound for "parents"); "husband" appears as *xediy*, *xediw*, and *xediyu*. Some of these variants may reflect mishearings; for instance, I know from my own experience how difficult it can be sometimes to distinguish final *i* from *iy*, or *esiyu* from *esyu*. But others are clear differences, as with the three-way opposition among the forms for "husband" just cited; in particular, the difference between forms in *-iw* and those in *-(i)yu* is quite distinct, and is recognized by Tsez speakers as a dialect or subdialect shibboleth. Nikolayev & Starostin (1994) reconstruct most of the relevant items as ending in **Vyu* in Proto-Tsezic, thus presupposing different simplifications in different Tsezic languages and dialects.² What is common to the variants is that they involve high vowels and glides, though typically without any gender distinction, i.e., a particular dialect will tend to choose a particular termination and apply this to all the relevant kin terms. In the case of Mokok *bab(i)yu* "father", it is possible that the dialect's favorite termination has even been added to a hypocoristic neologism, cf. Cebari vocative *baba* "father".³ The original function of this final **Vyu* is not clear. In their discussion of the word for "sibling", Nikolayev & Starostin (1994: 669) suggest that it may have marked female gender, but there is no evidence for this from any of the other kin terms, although the possibility certainly cannot be excluded — compare Russian, where final *-a* on nouns is basically a marker of feminine gender, but also occurs on a number of male kin terms, such as *djadja* "uncle", *papa* "Dad". What is clear is that in most of the present-day languages and dialects, the forms have some reflex of this **Vyu* without any distinction of gender. There are, however, two striking exceptions to this pattern, and both concern words for siblings.

In most of the Tsezic languages and dialects, there is a single word for "sibling", with no gender distinction. In the data presented in Table 1, however,

there are two varieties that deviate from this general pattern. In Bezhta, "brother" is *is*, while "sister" is *isi*. Significantly, the final vowel that marks the female form is the high front vowel *i*, the syllabic equivalent of the class II (female) agreement prefix *y-*. Even if, as suggested by Nikolayev & Starostin (1994: 669), this final *i* is a reflex of Proto-Tsezic **Vyu*, it is still the case that in Bezhta only this specifically feminine kin term has this termination. (Most of the others have *o*, corresponding to Tsez *u*; this includes, incidentally, the word for "mother", Bezhta *iyō*, cf. also Hinukh *iyō*, Hunzib *iyu*.) Thus the reason why this final *i* has been specifically associated with female gender is not fully explainable in terms of regular phonological developments. I would argue that it reflects the correlation between high front vowel/glides and female gender.

Even more striking is the distinction between *esiw* "brother" and *esiy* "sister" found in the Mokok dialect of Tsez.⁴ This is a remarkable distinction between this dialect and other dialects of the language, with a single word for "sibling". In general, Tsez dialects are very close to one another — with the exception of the aberrant Sagada dialect — with almost complete mutual intelligibility, and "translating" from one dialect to another involves in general only a few phonological and morphological adjustments, occasionally also changes in lexical items. But consider (9), which is the title of one of the Mokok dialect texts collected by Isa Abdullaev:

- (9) *ʔoʔno esiw-n sis esiy-n*
 seven brother-and one sister-and
 "seven brothers and one sister"

Already in the title, and consistently throughout the story, the different forms *esiw* and *esiy* are used to contrast brother and sister. The only way of "translating" this into another variety of Tsez would be by using a modifier that would specify the gender that is intended in each instance. But perhaps even more interesting is the way in which the distinction is made: female *y* versus male *w*. As with some of the other examples considered, it is not clear whether the differentiation is, in whole or in part, an archaism or an innovation. If it is an archaism, this may first entail revisions to the Proto-Tsezic forms suggested by Nikolayev & Starostin, but more importantly it would signal the tenacity of the correlation between *y* and female gender, and between *w* and male gender. The latter would be particularly remarkable, given that *w* is otherwise associated with male gender only in loans from Avar.⁵ If the discrimination is an innovation, then the male form in *w* can only reflect influence coming originally from Avar, as this is the only source of *w* as a class I correlate in the modern Tsezic languages. The female form could then either reflect that same influence, or the

continuing influence of the indigenous correlation between *y* and class II.

At this point, it is going to be useful to return to the noun class correlation of the high back vowel *u* in indigenous words (and potentially also of *w*, though I know of no relevant examples other than *esiw*). In Comrie & Polinsky (1999), it is argued that synchronically there is a correlation between both *i/y* and *u/w* and class II, and that this correlation is manifested particularly in inanimate nouns. The source of the correlation with *i/y* is historically transparent, as this correlation goes through a number of Nakh-Daghestanian languages and probably back to Proto-Nakh-Daghestanian. The source of the correlation between *u* and class II is less obvious, since the evidence of other Nakh-Daghestanian languages would suggest rather a correlation with class I. The best suggestion that I can make remains that made in Comrie & Polinsky (1999), namely that certain high-profile class II nouns precipitated this identification as an innovation, e.g., the word *baru* 'wife', which has this same form across all Tsez dialects for which I have information. Thus, while the correlation between *i/y* and class II seems to have been stable across the history of the Tsezic languages, the correlation between *u/w* and class I has been disrupted not only by the loss of any overt prefix for class I agreement,⁶ but also by the innovative correlation between *u* and class II; yet the original correlation has been given a new lease on life by the incorporation of Avar loans that explicitly manifest *w* as a class I marker.

In conclusion, the correlation between noun class, in particular as it reflects the male/female opposition, and the segments *i/y* and *u/w* in Tsez and the other Tsezic languages reflects an intricate interplay between synchronic and diachronic factors. The data available to us for these languages are still fragmentary, and are exclusively from the modern period, although Nikolayev & Starostin (1994) represents an important step in historical reconstruction. At present, it is not possible unequivocally to determine precisely which patterns reflect ancient phenomena, perhaps reinterpreted over the course of the history of the languages, and which are innovations, perhaps resulting from contact with the traditional lingua franca Avar. But at least it is now possible to formulate the questions clearly and to elaborate the interaction of general principles and language-specific data in teasing apart the diachronic and synchronic strands.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant SBR-9220219. I am grateful to Ramazan Rajabov, research assistant to this project, for his work on the Tsez dictionary (Cebari subdialect of Asaq dialect) (Rajabov MS) as part of this project, and to Madžid Xalilov for making available to me the partial manuscript of his Tsez dictionary (primarily

Mokok dialect) (Xalilov MS). I am also grateful to Isa Abdullaev for making available to me his collection of Tsez texts, especially those in Mokok dialect, to Arsen Abdulaev for his help in the interpretation of these texts, and to the Institute of Language, Literature and Art of the Daghestan Scientific Center of the Russian Academy of Sciences in Maxačkala for allowing me to consult the late Abdulkarim Abdullaev's unpublished candidate's dissertation (Abdullaev 1981). Finally, I am grateful to Maria Polinsky for our many discussions concerning Tsez.

Author's address

Bernard Comrie
Max Planck Institute for Evolutionary Anthropology
Inselstraße 22
D-04103 Leipzig
comrie@eva.mpg.de

Notes

1. This presumably relates only to some Asax subdialects. It is not, for instance, applicable to Ramazan Rajabov's usage (Cebari subdialect).
2. However, they also reconstruct a by-form **?obu* for "father", and suggest a distinction in Proto-Tsezic between **is*: "brother" and **is:Vyu* "sister", a point to which I return below. (Note that for consistency I have replaced Nikolayev & Starostin's *j* by *y*, for the high front glide.)
3. Nikolayev & Starostin (1994: 286), however, derive the Tsez form *babyu* from Proto-Tsezic **babVyu*, in turn from Proto-Nakh-Daghestanian **babayV*.
4. This may not be true of all subdialects of Mokok. For instance, Xalilov (MS), which is based on the Mokok dialect, gives only *esyu* for both "brother" and "sister". However, the differentiation of *esiw* from *esiy* by gender is consistent in a number of the Mokok texts collected by Isa Abdullaev.
5. It is, of course, possible that the female form in *y* reflects the indigenous correlation between class II and *y*, while the male form in *w* reflects Avar influence.
6. Some Tsez pronouns have a characteristic vowel *u* in some class I plural forms, e.g., *mežu-s* "our (class I)", *meža-s* "our (classes II–IV)", where *-s* is the genitive-1 suffix (used to modify a noun in the absolutive case). However, I doubt whether such forms are relevant. First, this use of *u* is restricted to the plural, and Tsez, like other Nakh-Daghestanian languages, has different class affixes in the plural from in the singular, namely Tsez IPL *b-*, II-IVPL *r-*. Second, the corresponding class II–IV vowel *a* is completely distinct from the class prefix, whether singular or plural. Third, there are some singular pronouns where *u* actually marks non-class I, e.g., *howda* "this (class I)", *howdu* "this (class II–IV)". Pronoun morphology involves a number of idiosyncrasies, and I see no way in which these idiosyncrasies throw any light on the affixes found as agreement affixes. It should be noted that Polinsky & Jackson (1999) also find evidence for diminishing relevance of *u* as a predictor of class assignment in Tsez.

References

- Abdullaev, Abdulkarim K. 1981. *Slovoobrazovanie v cezskom jazyke*. Candidate's dissertation. Maxačkala: Institut IJaL, Dagestanskij filial AN SSSR.
- Bokarev, Evgenij A. 1959. *Cezskie (didojskie) jazyki Dagestana*. Moscow: Izd. AN SSSR.
- Comrie, Bernard & Maria Polinsky. 1999. "Gender in a Historical Perspective: Radial categories meet language change". *Language Change and Typological Variation*, Vol. 2: *Grammatical Universals and Typology* ed. by Edgar C. Polomé & Carol Justus, 566–589. (Journal of Indo-European Studies Monographs, 31.)
- Imnajtšvili, David S. 1963. *Didojskij jazyk v sravnenii s ginuxskim i xvaršijskim jazykami*. Tbilisi: Izd. AN Gruzinskoj SSR.
- Nikolayev, Sergey L. & Sergey A. Starostin. 1994. *A North Caucasian Etymological Dictionary*. Moscow: Asterisk.
- Polinsky, Maria & Dan Jackson. 1999. "Noun Classes: Language change and learning". *Cognition and Function in Language* ed. by Barbara Fox, Dan Jurafsky & Laura Michaelis. Stanford: CSLI.
- Rajabov, Ramazan. In press. "Noun class in Tsez". *NSL-9* ed. by Howard I. Aronson. Chicago: Chicago Linguistic Society.
- . MS. *Tsez Dictionary*. Los Angeles: University of Southern California.
- van den Berg, Helma. 1995. *A Grammar of Hunzib (with texts and lexicon)*. Munich: Lincom Europa.
- Xalilov, Madžid Š. 1995. *Bežtinsko-russkij slovar'*. Maxačkala: Institut JaLI DNC RAN.
- . MS. *Cezsko-russkij slovar'*. Maxačkala: Institut JaLI DNC RAN.

CHAPTER 41

Once Again on the Reading of the Old Korean 尸

Alexander Vovin
University of Hawaii at Manoa

There is a considerable amount of controversy in the literature as to how to read the Old Korean phonogram 尸. The traditional point of view, presented by the majority of scholars in and outside Korea, holds that this phonogram has to be read as /l/. However, in 1956 this point of view was challenged by Yu Changkyun, who suggested that, on the basis of the Middle Chinese reading /s/ and internal Korean evidence, this sign has to be read as a fricative. Yu's proposal seems to be unaccepted by Korean scholars (cf. Yi Kimun [1961], who reads this character as a lateral, though he is aware of the problem [1961: 63]). In 1973 Yu Changkyun and Hashimoto Mantarô co-authored another article which suggested a certain compromise: in some cases 尸 had to be read as /l/, in others as /s/. This article relies heavily on the evidence from Middle and Old Chinese.

Roy A. Miller, obviously unaware of Yu Changkyun's work, came to conclusions virtually identical to Yu (1956), which he published in a series of articles (Miller 1979a, 1979b), suggesting that 尸 has to be read as /s/ in both the Koguryo language and Old Korean (Silla). Miller further sought to support his theory with external evidence from Altaic languages. As far as I know, his proposal is not supported by any other American historical linguist working in Korean (personal communication from Samuel E. Martin, S. Robert Ramsey, John B. Whitman, and J. Ross P. King). In Europe, however, Miller was supported by Sasse (1982), who tried to offer further evidence in favor of the fricative reading.

Yu Changkyun and Hashimoto Mantarô suggested that 尸 had a fossilized reading /l/ going back to the Old Chinese initial of that character which Hashimoto reconstructed as *sl- (Yu & Hashimoto 1973). Kim Wancin found this proposal unsatisfactory, pointing out the time gap between Old Korean and Koguryo on one side and Old Chinese on the other (Kim 1986: 29). Instead, he

came up with the hypothesis that the character 尸 is a result of an inadequate perception of the seal script form of the character 乙. My feeling is that his theory is a little bit forced: while Kim Wancin explains other character replacements by the similarity of their cursive script forms, here he has to appeal to evidence from the seal script, since the cursive forms of 尸 and 乙 are quite different.

The goal of the present study is to continue the line of Yu Changkyun and Hashimoto Mantarō by providing additional evidence that 尸 has to be read as /l/ and that this reading is based on Old Chinese. I will examine different kinds of evidence supporting the above claim, and refuting the Miller-Sasse position. The main evidence comes both from Silla and Koguryo writing systems as well as from the Old Chinese reconstruction.

(I) If in the early post-Karlgren reconstructions of OC,¹ the initial of the character 尸 was reconstructed as *sl- (although J. Norman already suggested voiceless *lh- in 1969 [according to Yu (1994: 163)]), which certainly leaves room for doubts as to whether such a complex initial could be used to write [l] or [r] in Old Korean, then in the up-to-date Chinese reconstructions, the Old Chinese initial of the character 尸 is reconstructed as *hl-, not as *sl-, which is quite crucial in this case (Baxter 1992: 787).

(II) First, I intend to examine the use of Old Chinese readings in Koguryo fragments. The character 尸 appears in phonetic glosses only twice: in the word 也尸 “wolf”, which Miller compares with the MK *yezo* “fox” (see below), and in the word 買尸 “garlic”, compared by Miller with OJ *mira* “id.” and WM *manggirsun* “id.”. The Old Chinese readings of the characters 也尸 are *ljaj-hljij. It is really tempting to compare the corresponding Middle Chinese reading *ya-ši with the MK *yezo* “fox”, as Miller did, but this resemblance turns out to be an illusion. First, as Yi Kimun demonstrated, MK /ye/ almost always goes back to an earlier *i (Yi 1959). Second, as the author of these lines suggested, the MK grapheme Δ “triangle” might have been in fact pronounced /ñ/, and not /z/, as usually suggested (Vovin 1993). Thus, we arrive at the tentative Old Korean form *iño “fox”, which does not resemble *ya-ši at all. Given in addition the semantic difference, Miller’s comparison can be safely dismissed. It is possible, however, that *ljaj-hljij transcribes some cognate of the Middle Korean (MK) *ilhi*, *ilhi* “wolf” (Yu 1986: 625; Cahwoy: Sang 10r),² also attested as modern standard Korean *ili* and in dialects: *ili* (Kyengsang, Chwungpok, Kangwen), *ilittay* (Celnam) (Choy 1987: 937). Since initial *n-* in front of /i/ or /y/ in MK was already unstable (cf. *ilum* “name” vs. *nilo-/nilu-* “to name”, “to call”), it is quite possible that *ljaj-hljij transcribes something like *niähli. There are, of course, arguable points in this etymology, but they are not unsurpassable, and do

not require us to make a “fox” out of a “wolf”. Moreover, it is obvious that a comparison of 也尸 “wolf” with MK *ilhi* is possible only if we assume that the character 尸 was used according to its phonetic value in OC and not MC.

An even more supportive example involves the Koguryo word 買尸 “garlic” which Miller reads as **mais/*maís* comparing it at the same time with OJ *mira* “id.” and WM *manggirsun* “id.”. If this comparison is true (at least its OJ part), then it is unclear why OJ exhibits *-r-* and not *-s-*: by Miller’s own set of correspondences reflecting PA **-l₂-*, OJ must have *-s-* (Miller 1971: 115). Noting this first inconsistency, we can as well note another: there is a very strange correspondence of Koguryo *-ai-* or *-ei-* (if we follow Miller in reading the character 買 according to its MC reading [mei]) to OJ *-i-*. Miller is not alone, however, in reading this character according to its MC reading. Another famous example among Koguryo glosses is the word “water” or “river” also written as 買. It is customary to compare this Koguryo word, which we should read as [mei] if we use the MC reading of the character, with OJ *midu* “water” or MK *múl* “id.” (Yi 1961). What attracts one’s attention again is the bizarre vocalism in the Koguryo: it is difficult indeed to demonstrate how Koguryo vocalism can be coordinate with OJ or MK vocalism. However, there is no problem if we assume that this character was used according to its OC reading **mri*. It has been suggested that the Proto-Altaic word for “water” is to be reconstructed as **mür̥i* (Starostin 1991: 278). OJ *midu* and *mina* “water” is likely to go back to **mi-*, which in its turn may go back to pre-PJ **miri* according to Whitman’s law (Whitman 1985). This pre-PJ reconstruction **miri* remained hypothetical until now, since there is no internal Japanese basis to demonstrate *r*-loss. The Koguryo evidence, if treated appropriately through the prism of the OC phonology, demonstrates unambiguously that there was indeed *r*-loss. The Koguryo form can be explained as a result of the development **mür̥i* > **mri*, with a reduction of the vowel of the first syllable.³ Then the Koguryo word 買尸 “garlic” should be read as **mrihljij* according to the OC readings of the characters, denoting in all probability something like **mür̥(i)*, which is in its turn a likely development from still earlier **mür̥(i)*. Recalling that the suggested parallel to this word is OJ *mira* “garlic”, we discover again that using OC readings minimizes our problems with the vocalism. MK *mànól* “garlic” must be related as well, since we can see here the same development as described by Whitman in the case of OJ *pari* “needle” (< **parari* < **panari*) and MK *panól* “id.” (Whitman 1985).⁴ I believe that WM *manggirsun* is not related to a tentative PJK **mün̥ir*, since WM *-ngg-* does not correspond regularly to PJK **-n-* although it may be connected to it through borrowing. Thus we can see that Koguryo fragments indeed are based at least partially on OC readings, and, therefore, Kim Wancin’s argument that OC

is separated by a significant time gap from Koguryo and Silla has to be abandoned in the light of the evidence presented above.

(III) The Old Korean writing system is not exclusively based on MC either, but includes certain elements of OC, as Yu & Hashimoto have already suggested. Thus, they demonstrated that the character 支 in Hyangka has to be read according to its OC reading *gi (*kje in the more up-to-date reconstruction by Baxter 1992), and not the MC reading *tsi (Yu & Hashimoto 1973: 7–10). Let us have a look at another example: the reading of the character 邪 in the Hyangka IV–5:

耆郎矣白史是史蔽邪

Kim Wancin deciphers this line as: *Kilang-oy cuz-i wolsi swuphuliya*, and translates it into modern Korean as *Kilang-uy mosup-i olsi swuphwul-iyē*, meaning “[Hwa]lang Ki[pha]’s appearance is right being [like] a forest” (Kim 1986: 81, 91). Syntactically, though it is strange to have either the gerund marker *-a* or question marker *-ya* in this sentence: we should expect a final verb here. Therefore, I believe, Yu Changkyun is right to reconstruct this line as: *Kul molo-uy cus-i isiswola* “There is appearance of hwarang Kul[pha]” (Yu 1994: 407) on the basis that the OC reading of the character 邪 is *lja (*ra in the reconstruction used by Yu). Since the character 邪 is read as *ja or *za in MC (resulting in Sino-Korean readings *ya* and *sa* respectively), it cannot be deciphered appropriately by Kim Wancin, who built his study of Hyangka essentially on Sino-Korean readings of the characters. Examples of this kind are abundant in Hyangka, and the recent study of Yu Cangkyun (1994) provides many interesting examples.

It is certainly out of place here to survey all instances of 尸 usage in Hyangka texts: it takes three charts inserted in the concordance to Hyanka compiled by Sasse (1989: 65–67). However, in order to see that 尸 can be interpreted as a lateral and there is no need for a fricative, I will give several examples:

- (1) 哭屋尸以憂音 (I: 2)
 WUL-wo-I-LO SILUm
 sob-?-PART-INS sadness
 “sadness [spent] in sobbing”

Here we have two characters, 屋尸, denoting some kind of verbal form, most likely corresponding to the MK irrealis participle in *-lʔ*, plus some suffix *-wo-* preceding it (modulated form?). The explanation is not unproblematic: first, the

modulator after a [-back] stem is expected to be *-wu-*, not *-wo-*; second, there seem to be no cases when the MK participial form in *-lʔ* is followed by the instrumentalis *-(o/u)lo*. Nevertheless, this solution seems to be better than setting up a completely new verbal form in *-s* unparalleled by MK data.

- (2) 慕理尸心未行乎尸道尸 (I: 7)
 KULi-I MOSOm-oy NI-wo-I KIL
 yearn-PART heart-GEN go-?-PART way
 "the way that [my] yearning heart goes"

Here we again have 尸 twice corresponding to the MK irrealis participle in *-lʔ*, but contrary to the previous example, this one is completely unproblematic. In addition, we also have the word *kil* (MK *kilh*) with the last sound written phonetically.

- (3) 二尸掌音 (VII: 2)
 TWUPUI SWONPATAm
 "two handpalms"

Here 尸 can be equated with *-l* of MK *:twulh*, EMK *twubwul* (attested in Kyelim Yusa, #20)⁵ "two".

In all these cases, Korean internal evidence demonstrates a lateral in all historical materials and in all dialects. The fricative hypothesis cannot be saved by appeal to Altaic data, either. The 尸 as an irrealis participle marker (MK *-lʔ*) in all likelihood corresponds to the Manchu imperfective participle marker *-ra/-rel-ro*, the OJ participle marker *-uru*, and possibly the Turkic aorist marker *-Vr*. Using the correspondence Korean *l*: Manchu *r*: Japanese *r*: Turkic *r*, it is possible to reconstruct only PA **r₁*, and not a voiceless fricative. The word 二尸 TWUPUI "two" (MK *:twulh*, EMK *twubul*) corresponds to Manchu *juwe*, Ewenki *ju:r* "two", and Middle Mongolian *jirin* "two (of females only)", which again point to PA **r₁*,⁶ but not to a voiceless fricative. There are also regularity problems with the Altaic evidence presented for the fricative nature of 尸, to which I will turn next.

(IV) I also believe that Miller's Altaic evidence for /s/ is not tenable because Miller's correspondences are chaotic: his Old Korean /s/ corresponds not only to Altaic **l₂*, which is expected, but to Altaic **r₂* as well, which is bizarre, given the reflexes of **r₂* in the Altaic daughter languages:

PA	PJ	PK	PMT	PM	PT
*l ₂	*s	*l	*l	*l	*s
*r ₂	*t/*ɾ	*l	*ɾ	*ɾ	*z

For example, Miller compares MK *:solp-* “to inform” with OK 白屋尸 “id.”, which he reads as *sius-/*sius- and compares with PA *ser₂- “to know” (Miller 1979a: 361). Right before this example we find another comparison, involving both MK *kél* “reed” and MK *:kwol* “rush, sedge” as well as OK 古尸 “reeds”, which Miller reads as *kos/*koś and compares with Middle Turkic *kaš-ak* “reed, sedge” (< PA *ka₂) (Miller 1979a: 361). This inconsistency, of course, does not contribute either to Altaic theory or to progress in the study of the Korean writing system, but rather obscures both matters. There is no controversy if we read it as [l], which is an expected reflex in Korean of both PA *r₂ and *l₂. This inconsistency was further aggravated by the fact that Miller distinguished two groups of examples including 尸: A, in which MK has the fricative counterpart to the 尸, as in 也尸 “wolf”; and B, where MK has /l/ corresponding to 尸, as in 買尸 “garlic”. This gives us at least two possible MK correspondences and three PA correspondences to 尸:

OK and Koguryo	Middle Korean	Proto-Altaic
尸	s, z	*s
尸	l	*l ₂
尸	l	*r ₂

Finally, OK 尸 as evidence for a special reflex of PA *l₂ in Korean becomes almost worthless as it can correspond to three PA phonemes.

(V) As mentioned previously, Sasse 1982 tried to give additional evidence for the fricative reading of 尸. His strategy was somehow different from Miller’s as he mostly concentrated on Miller’s group A, trying to increase the number of cases where 尸 corresponds to a fricative in the later stages of Korean. I will survey below all Sasse’s examples from group A:

(1) Old Paekche 古尸, Middle Korean *kwós*, Modern Korean *koc* “cape” (Sasse 1982: 713). There are several problems: first, neither Yu 1964 nor Nam 1960 include Middle Korean *kwós* “cape”, nor it is attested in *Hwunmong cahwoy* 1527, so Sasse’s source for the Middle Korean attestation remains a mystery to me. Second, modern Korean has *koc*, with a final affricate, and not a fricative.

The other dialect attestations also indicate only /-c-/ and no /-s-/ (Choy 1978: 64). Therefore, we will get a fourth source for our OK 尸. Third, neither Middle Korean nor Modern Korean are direct descendants of the Paekche language. Finally, if one equates 古尸 with 岬 in the phrase 岬城郡本白濟古尸伊渠 (Samguk Sagi 36:3), it is problematic to find a correspondence to 城. It is unlikely to be just 伊, which is in all probability read as [i], while the MK word for “castle”, “fortress” is :cas. Sasse also cites the following passage from the “Samkwuk Yusa” (IV.183:10):

嘉瑟岬。或作加西。又嘉栖。皆方言也。岬俗云古尸。·有加西岬。或云嘉瑟岬。

“*Kapi* — ‘cape’. Some say *kasi*. Also *kasi*?. They are all dialectisms. Vulgar name for ‘cape’ is *kol*. ... there is mountain pass *kasi*, some say *kapi*” (translation is mine — A.V.) where the word *cape* is written as 嘉瑟, 加西, 嘉栖 which are all said to be dialect variants of the word for “cape”. Let us note that the character 瑟, while it has a MC reading *srit, was read as *sprjit in OC, which is demonstrated by its phonetic element 必 *prjit. Therefore, I believe that it reflects *kapi “cape”, attested in the Swunchen, Koyyang, and Cinto regions of Southern Cenla province as *kapi* or *kaypi* (Choy 1978: 64).⁷ The other two words, 加西 and 嘉栖, are to be read as *kasi* and *kasi*?. (西: MC *sej and OC *sij; 栖: MC *sejX and OC *sij?), so there is no doubt that they have a fricative -s-. However, it seems difficult to equate them with Korean *koc* or Paekche 古尸, as they exhibit different vocalism, and it is safer to assume that we are dealing here with three unrelated words. Thus, we can see that the above passage from the “Samkwuk Yusa” provides almost no support for Sasse’s claim.

(2) Koguryo 波尸 “barnyard millet, panic grass” is compared by Sasse to MK *pyé* “rice plant”, which, as Sasse argues on the basis of the Chinese transcription 別思 attested in the “Cosen kwan yek.e” (after 1400 C.E.), once had a final fricative -s (Sasse 1982: 713–14). Sasse’s phonetic reconstruction is persuasive, but it remains unclear what “barnyard millet” has to do with “rice plant”. In addition, as demonstrated by Yi Kimun, MK *ye goes back to PK *i, so in this case we again are faced with highly divergent vocalism: *pal vs. *pis. In combination with doubtful semantics, this comparison is better abandoned as well. It seems to me that Koguryo *pal might have a better etymology in OJ *apa* “millet”, if we assume that the Koguryo form underwent apheresis.

(3) Sasse also attempts to equate 尸 the with Middle Korean genitive marker -s (Sasse 1982: 714–15). Even leaving aside a problem, that this genitive marker may in fact go back to an archetype *-ñ as demonstrated in Vovin 1993, there are problems with Sasse’s examples:

- (3) a. 新良県本百済沙尸良県 (Saki 36: 3)
 b. 大山郡本百済大尸山郡 (Saki 36: 3)
 c. 帶山県大尸山 (Saki 37: 3)

In example (3a), Sasse equates 沙尸 with MK *sáy* “new” plus the genitive marker *-s*. Given MC **sræ* and OC **sCraj* for 沙, as well as the fact that MK *sáy* is likely < *sa[l]i, that may not seem to be unreasonable at first glance, but as a matter of fact the whole proposal is based on two false assumptions. First, Sasse suggests that MK *sáy* means both “east” and “new” (Sasse 1982: 714). This is not true, as MK *sáy* means only “east” (Yu 1987: 429; Nam 1987: 296),⁸ while MK *sáy* means only “new” (Yu 1987: 441; Nam 1987: 295). In 15th C. texts MK *sáy* “east” and MK *sáy* “new” have different spelling (MK *sáy* “east” is attested in the “Kumkang kyeng sam ka hay” (1482 C.E.) and MK *sáy* “new” is attested in “Welin sekpo” (1459 C.E.)) and, therefore, have a different etymological origin. The confusion started considerably later, when the MK vowel /o/ disappeared from the system, so we have unetymological spelling *soy* “new” in the “Chengkwa yengen” (1728 C.E.) (Yu 1987: 429). Based on this false equation, Sasse then provides an example of how the genitive marker *-s* is used after his alleged *sáy* “east, new”: *sáy-s palam* “east wind”. But this of course is applicable only to MK *sáy* “east”, not the MK *sáy* “new”, which always appears in MK texts without the following genitive marker:

sáy kwùsúl-i nà-myé (Welin 1: 27)

new treasure-NOM appear-GER

“new treasure appears, and...”

sáy cwóh-òn wós nìp-kwó (Nung 7: 6)

new clean-PART clothes put on-GER

“having put on new and clean clothes...”

Sasse believes that in example (3c) both the character 帶 and the character 大 are phonetic spellings of the word *tay* “bamboo” (MK *tay*). He comes to this conclusion on the basis of the following gloss:

帶方州本竹軍城 (Saki 37: 4).

His speculation, however, goes against the fact that in geographical glosses of the Samkwuk Saki, one part tends to use characters as logograms, while the other as phonograms. If it is true in the above example, why should it be

different in example (3c)? It seems completely unreasonable that in the geographical records of the Samkwuk Saki, which had a goal of explaining meanings of obscure place names, or to give an old phonetic spelling of the newly adopted logographic spelling, the word “bamboo” would be written twice with different phonograms as if with the purpose of obscuring the matter further. Moreover, example (3b) seems to be very transparent: the character 尸 in the combination 大尸 seems to be used in the same way as it is used in Hyangga: in the capacity of a phonetic key to a logogram. It is likely that we have just an equivalent of the MK participle *-lʔ* here, and the sequence has to be read as *ha-l* “big-PART”. One can disagree, saying that in MK “big mountain” is more likely to be expressed as *ha-n mwoy*, with a realis participle *-n*, instead of irrealis *-lʔ*. However, we should not forget that this is a specimen of the Paekche language, and we do not know exactly whether participial functions were one hundred percent identical in the Paekche language and Middle Korean. In short, I believe it is possible to reject Sasse’s examples (3a), (3b), and (3c) as evidence for the fricative nature of 尸.

(4) The example of a king’s name written as 伊品 (Yusa II: 118.7) or 伊尸品 (Yusa II: 118.8) is of little value for proving the sibilant nature of 尸 here, since Sasse himself notes that “as there are no other clues concerning this name, it is difficult to give an explanation”, so Sasse’s speculation that it may reflect genitive *-s* (Sasse 1982: 715) remains, I am afraid, a speculation unsubstantiated by any evidence.

(5) The last example adduced by Sasse for proving the fricative nature of *-s* is of paramount importance for proving that this hypothesis is incorrect. Here is the passage Sasse cites:

... 我名未尸 (Yusa III: 154.8)

“... my name [is] Mil.”

... 未與彌聲相近。尸與力形相類。 (Yusa III: 154.11)

“... tones of [character] 未 and [character] 彌 are mutually close, forms of [character] 尸 and [character] 力 are mutually similar.”

Sasse goes on to argue that while 未 (MC *mjijH, OC *mjijS) and 彌 (MC *mjieX, OC *mjejʔ) are really similar in sound values, 尸 (MC *syij, OC *hljij) and 力 (MC *lik, OC *C-rjik) are not, and he emphasizes that while the text mentions that the first pair is similar in sound, it refers to the similarity of the form for the second pair (Sasse 1982: 716). First of all, of course, the graphic form of 尸 and 力 is not similar: it is quite obvious. It is not similar in cursive writing either. Second, it is of course unreasonable to require Ilyen in the 13th C.

to use the phonological terms of the 20th. Let me note that he is not accurate with the terms of contemporary Chinese phonology. Thus, the term 声, which is used to describe the similarity of the first pair of characters, refers in Chinese phonology not to syllable (and certainly not to sound, the notion of which was completely unknown to Chinese phonologists, although Sasse seems to believe otherwise, translating the passage as “similar in sound”), but to tone. It is clear, though, that the tone difference is not what Ilyen meant: 未 has a departing tone, and 彌 has an rising tone. Ilyen, of course, refers just to phonetic similarity, in the case of the first pair as well as in the second, since we know that 尸 has an OC reading *hljij with an /l/. Therefore, we have an important witness here: Ilyen himself, who testifies that in the OK writing system 尸 was used phonetically in the same way as 力: to write a lateral /l/, not a voiceless fricative /s/ or /ʃ/.

(6) Sasse has also sought to provide internal Korean support for yet another of Miller’s reconstructions with a fricative, namely Koguryo *kunš “writing”, reconstructed on the basis of 斤尸 in the following passage:

文峰県一云斤尸波兮 (Saki 37:4)

Miller compared this with PA *kal₂- “to scratch”, “to dig”, which he also compares to yet another Koguryo word 加尸 “plough”, which Miller reconstructs as *kaš (Miller 1979b: 13–14). This seems impossible, as it would require PA *kal₂- to split into two completely different etymons in Korean, violating in this way the regularity of other correspondences. Sasse, while certainly aware of the existence of MK *kùl*, *kùlwál*, *kùlwél* “writing”, compares Koguryo *kunš with Korean *kus*- “strike a match, mark (draw) a line, write (a Chinese character or a stroke of Chinese character)”, which he claims to be a perfect semantic and phonological fit (Sasse 1982: 712). I am not sure, though, that the fit is so perfect. First, none of the meanings directly connected with writing cited by Sasse for Korean *kus*- are attested in the dictionaries: there one can find only “draws”, “strike a match, mark (draw) a line” (Martin 1967: 257; Han 1994: 212). Of course, dictionaries often don’t tell the whole story, but this at least demonstrates that this meaning, even if does exist, is marginal. Moreover, all the numerous examples of MK *kùz*-, given in Yu (1987: 107) have only the meaning “to drag” (modern Korean *kkul*-). Finally, MK *kùz*- may be actually *kùñ*-, as suggested in Vovin 1993.

To sum up, Sasse has built a more thorough case than Miller, but I believe that all his examples in favor of the fricative nature of 尸 are not valid. The character 尸 in Old Korean and Koguryo writing stood for lateral /l/ (or rhotic /r/), being one of the leftovers of Old Chinese in the system which mostly utilized Middle Chinese readings.

Texts

Kyelim	Kyelim Yusa, 1102
Kumkang	Kumkang kyeng sam ka hay, 1482
Nung	Nungam kyeng enhay, 1462
Saki	Samkwuk Saki, 1147(?)
Welin	Welin Sekpo, 1459
Yusa	Samkwuk Yusa, 1285

Author's address

Alexander Vovin
 East Asian Languages & Literatures
 University of Hawaii at Manoa
 1890 East West Road
 Honolulu, HI 96822, USA
 vovin@hawaii.edu

Notes

1. In this article the following abbreviations are used: EMK — Early Middle Korean, MC — Middle Chinese, MK — Middle Korean, OC — Old Chinese, OJ — Old Japanese, OK — Old Korean, PA — Proto-Altaic, WM — Written Mongolian.
2. Yu (1994: 156) also suggests that this Koguryo word corresponds to K *ili*, though he does not mention the MK form and essentially offers no explanation.
3. A parallel case is offered by the Koguryo numeral “three”, written by the character 密 “secret”. Its OC reading is *mrjit, suggesting the same *r*-loss in OJ *mi*- “three”, which is believed to be its cognate. As suggested in Vovin (1993: 256) and (1994: 247), MK *seys* R “three” < PK *ñe[C]i may also be related.
4. Cf. also OJ *mura* “village” and MK *mòñòlh* “id.”, suggested in Vovin (1993: 257), which also demonstrate the same development.
5. Numeration of entries in the *Kyelim Yusa* follows Kang 1991.
6. Since Turkic data are lacking, it is not clear whether one has to reconstruct PA *r₁ or *r₂.
7. In all likelihood, a loan from Sino-Korean *kap* (岨) “cape”.
8. Yu Changton gives MK *sòy* with low pitch, but Nam Kwangwu gives MK *sóy* with high pitch. They both refer to the same passage in the same text: Kumkang kyeng sam ka hay 2:6 (1482 C.E.). According to the fascimile edition I have (published by Hankul hakhyo chwulphanpu), the word indeed has high pitch: *sóy*.

References

- Baxter, William H. 1992. *A Handbook of Old Chinese Phonology*. Berlin & New York: Mouton de Gruyter.
- Choy, Hakkun. 1978. *Hankwuk pangen sacen*. Seoul: Hyenmunsa.
- Han, Kapswu, ed. 1994. *Kwuke tay sacen*. Seoul: Kyoyuk sekwan.
- Kang, Sinhang. 1991. *Kyelim yusa Kolye pangen yenkwu*. Sengkyunkwan tayhak chwulphanpu.
- Kim, Wancin. 1986. "Decipherment of Hyangga and Textual Criticism". *Korean Linguistics* 4.27-31.
- Martin, Samuel E. 1992. *A Reference Grammar of Korean*. Rutland & Tokyo: Charles E. Tuttle.
- , Yang Ha Lee & Sung Un Chang. 1967. *A Korean-English Dictionary*. New Haven & London: Yale University Press.
- Miller, Roy A. 1971. *Japanese and the Other Altaic languages*. Chicago & London: University of Chicago Press.
- . 1979a. "Old Japanese and Koguryo Fragments: A re-survey". *Explorations in Linguistics, Papers in honor of Kazuko Inoue* ed. by George Bedell et al., 348-368. Tokyo: Kenkyusha.
- . 1979b. "Old Korean and Altaic". *Ural-Altaische Jahrbücher* 51.1-54.
- Nam, Kwangwu. 1960. *Koe sacen*. Seoul: Ilcokak.
- Sasse, Werner. 1982. "F as a Phonogram in Early Korean Writing". *Linguistics in the Morning Calm. Selected Papers from SICOL-1981* ed. by Linguistic Society of Korea. Seoul: Hanshin.
- . 1988-89. *Studien zu Entzifferung der Schrift altkoreanischer Dichtung. Bd. I: Theorie und Praxis der Entzifferung (1988); Bd. II: Konkordanz Teil 1: Silla Hyangga (1989)*. Wiesbaden: Otto Harrassowitz.
- Yi, Kimun. 1959. "On the breaking of *i in Korean". *Aseyya yenkwu* 2:2.131-138.
- . 1961. *Kwuke sa kaysel*. Seoul: Thap chwulphansa.
- Yu, Changkyun. 1956. "Hyangka ey nathanan F munpepcek kinung kwa umka". *Kwuke kwukmunhak* 15.36-61.
- & Mantarô Hashimoto. 1973. "Hyangka phyoki yongca uy sangkosengcek chukmyen". *Sinla Kaya Munhwa* 5.1-29.
- . 1994. *Hyangka pihay*. Seoul: Hyengsel chwulphansa.
- Yu, Changton. 1964. *Yicoe sacen*. Seoul: Yensey tayhakkyo chwulphansa.
- Vovin, Alexander. 1993. "About the Phonetic Value of the Middle Korean Grapheme Δ". *Bulletin of the School of Oriental and African Studies*, vol. LVI.2.247-59.
- . 1994. "Genetic Affiliation of Japanese and Methodology of Linguistic Comparison". *Journal de la Société Finno-Ougrienne* 85.241-256.
- Whitman, John B. 1985. *The Phonological Basis for the Comparison of Japanese and Korean*. Harvard University Ph.D. dissertation. Cambridge, Mass.: Harvard University.

CHAPTER 42

The Role of Historiography in Evaluating the Results of Comparative Linguistic Work

A case study

Paul Sidwell

Australian National University

It is mainly Koerner's work which has served as the bridge from barren, self-serving accounts to a professional historiography which deals with the history of linguistics in its wider context. Koerner's extraordinarily profound knowledge of primary sources, his ability to apply sound methodology to the linguistic and historical situation of a given milieu, his ardent and unshakeable belief in the victory of the cause of truth, and his outstanding talent as a writer have made Koerner the acknowledged leader of linguistic historiography. Koerner writes clearly and to the point — what is bad he directly calls bad, disdaining to conceal his evaluation behind the hedge of vague hints and innuendos. The fearless honesty of the Koernerian approach is as sharp as Occam's Razor, and as essential in the struggle for truth. We hold that no phenomenon in nature, including in languages, can be understood if taken by itself, isolated from surrounding phenomena; and that, vice versa, any phenomenon *can be understood and explained* if considered in its inseparable connection with surrounding phenomena.

Kirk & Sidwell (1998: v) "Professing Koernerian Linguistics"

1. Introduction

The present paper looks at the importance of the historiographic aspect of examining the results of Comparative Linguistics in the Austroasiatic¹ context. The investigation of the history of Austroasiatic languages is underdeveloped compared to the study of Indo-European, lagging by perhaps a century. However, this perceived "lag" should decrease rapidly as studies in Austroasiatic linguistics

progress — after all, the logic of the comparative method and the practical methodology has already been highly developed. On the other hand the early Indo-Europeanists were trailblazers who had to forge their own armory of principles and techniques. Assuming that this collection of principles and techniques reflects a universally applicable method in investigating linguistic history, and not just one which reflects unique facts of the Indo-European family, application to “new” language families should yield reliable results, and in less time than elapsed between William Jones’ famous speech and the publication of Pokorny’s dictionary.

In recent decades reconstructions of various Austroasiatic proto-languages have appeared in print, but how are we to assess their results? It is not just that the often preliminary nature of results can make assessment difficult; there is also a social factor. There are relatively few scholars in Austroasiatic historical linguistics, and individual languages or even linguistic groupings tend to be the domain of individual specialists, sometimes protecting their domains like oriental warlords. In other cases the specialists in a given language family will tend to be small groups of close colleagues, sharing common views and intentions. It is not like Indo-European studies, where one has little trouble finding scholars who are suitable to offer competent and independent assessment of a given proposal, even if they are not specialists in the particular language(s). In a relatively young field like Austroasiatic it becomes imperative for scholars who are somewhat removed from the immediate milieu to critically evaluate linguistic claims, particularly reconstructions. We argue that it is not enough to simply discuss the linguistic facts of an issue when discussing the merits of something as complex as a reconstruction. To understand why particular decisions were made, such as phoneme assignments or choice of languages, one must be aware of details of the immediate milieu within which the paper is written, and the historical conditions that gave rise to that context, in other words we need a comprehensive historiography, and one that takes the “human factor” into fuller account.

Unfortunately there are still quarters where such discussion is not welcome. It is very much an important aspect of the human-factor that personal insecurity and small mindedness leads individuals to take things personally without considering the greater duty to the discipline.² Of course we hear the often repeated principle that linguistic claims are only properly assessed on linguistic facts. While this is internally logical, it ignores that fact that linguistic inquiry is a human pursuit, and understanding why we do the things we do means looking beyond the narrow view of immediate circumstance.

Below we briefly review three Bahnaric reconstructions which were done in the 1960s, and which have since been treated largely as received wisdom. The

publications in question are presented as works of comparative reconstruction, and have many of the features one expects to see in such work. However, close examination shows some problems with their methodology which call into question the validity of their results. The nature of the problems and why they occurred can be explained by the underlying motivations and consequent approaches of the milieu from which this work emerged. Once these factors are taken into account we believe that considerable progress in Bahnaric reconstruction is possible without new or improved language descriptions.

2. The birth of comparative Bahnaric studies

From the late 1950s to the early 1970s, in the special circumstances created by the military-political situation in Indo-China, a generation of scholars associated with the Summer Institute of Linguistics, 'The Linguistic Circle of Saigon', conducted extensive fieldwork, collecting copious wordlists and texts, and publishing useful grammars of previously poorly studied languages. Among these young and highly motivated scholars some took a particular interest in historical phonology and reconstruction, and by their hands at least three reconstructions appeared in print, all prepared at approximately the same time. They worked independently of the earlier French tradition, and thus had to forge their own conventions and priorities. These latter day *Junggrammatiker* collected their own vocabularies and grammars in the field, sharing data and collaborating closely with each other in their analyses. Their training was typical of the 'Bloomfieldian' approach, common among American linguists into the 1960s, which emphasized recording of real utterances in the field, and they made phonemic analyses according to the methods and principles laid down by Pike (1947) and his followers. This was the dawn of a golden age for linguistics of Southeast Asia — previously fieldwork was much more scattered and unorganized, with far less collaboration between workers in the field. It was an era of tremendous excitement for these earnest young scholars. They were blazing trails through (linguistically) virgin territory, often in extraordinary personal danger. They had excellent training in descriptive linguistics, and their personal convictions motivated a genuine interest in finding the common ancestor of these closely related languages. Thomas (1964: 162) writes at the time (his excitement and optimism showing):

The lack of careful phonemic (or prosodic) descriptions has been the main hindrance to progress in Mon-Khmer linguistics and Austroasiatic comparisons. But fortunately this lack is now being rapidly filled from several quarters, so the next few years should see rapid progress.

The new generation revolutionized the field, and within a few years dozens of wordlists and other papers began appearing, with data generally collected according to standardized formats with transcription systems based upon Vietnamese orthography. This meant that there was now a qualitatively and quantitatively improved situation compared to that faced by comparativists such as Schmidt and Pinnow. The newcomers collected data on many previously undescribed languages, and devised novel solutions to the problems of working with them. They advanced our understanding of the phonetics, phonology, and word structure of these languages, and they made real advances in investigating the linguistic history.

However, our heroes formed a close scholarly community, with its own distinct attitudes and perspectives, and we should be mindful that their interest in historical linguistics was very much a function of their interest in bringing biblical scripture to the local peoples — in a form which would be accessible to as many speakers as possible. In such a milieu the idea of finding a somehow basic form of language which unites disparate dialects becomes a powerful motivation. With some decades of experience now, the SIL has a highly elaborated set of procedures for surveying closely related languages, systematically comparing them to reveal the common elements at the various linguistic levels, and then producing a standardized *Lingua Franca* which has a high degree of intelligibility to the speakers of the related tongues (see handbooks such as Grimes 1995). This type of work is what is taught as comparative linguistics among SIL scholars, so we need to be very mindful of this when reviewing their comparative work.

The three works of the SIL tradition which we review below were all prepared within two years (or less) of each other. Work on Blood's MA thesis could not have started before 1963, because of the sources he used, and was submitted in 1966. Thomas & Smith do not give any dates for their sources, but the paper appeared in print in 1967, and includes a reference to a paper in *Mon-Khmer Studies II* (1966), so that paper was probably written over late 1965 and/or early 1966. Smith (1972) is based on a manuscript, dated 1967, which the present writer has seen archived at SIL Bangkok. We mention these points because we wish to highlight the closeness of the circumstances within which these papers were written.

3. Blood (1966): A reconstruction of Proto-Mnong

Blood presents a low-level reconstruction based upon a comparison of Mnong dialects. The main purpose, Blood explains, is to test whether the historical

development of the vowels in a Mon-Khmer grouping can be successfully reconstructed. Blood draws particular attention to 'the vowel problem',³ repeatedly stating that comparativists have found it impossible to find regular correspondences among vowels of Mon-Khmer languages, and even asserting, without citing sources, that "Some linguists have even made the statement that regular sound laws do not exist for these vowel shifts, and indeed no published work to date has been able to refute this claim" (p. 1). Blood mentions various works which failed to find any regular patterns of correspondence among vowels, and decides to take up the challenge made by Thomas (no text reference is given by Blood for the quote):

I feel that the solution... lies in starting at the very lowest level of comparison, working on adjacent languages to establish proto-forms at that level, then using these reconstructions as the basis for comparison at the next level. Only in this way, I feel, will the Mon-Khmer vowels be able to be solved. (p. 7)

We feel, however, that the fuss made over 'the vowel problem' is peripheral to the real importance of Blood's work. Given the mixed quality of sources available until that time, the unfamiliarity of comparativists generally with languages of the region, and the lack of good classifications, it is entirely understandable that such difficulties would be encountered. This is a much more reasonable explanation than an invocation that the uniformitarian principle ceases to function in respect to various languages (such that sound laws do not exist). Blood's reconstruction concerns a group of dialects that are so closely related that the correspondences are rarely more than trivial in nature, so that correspondence is largely just a function of resemblance. Ultimately Blood's Proto-Mnong is essentially the same as his Eastern Mnong dialect, being the most phonologically conservative of the chain — we can characterize this as 'reconstructing a language from its dialects'. While this may not advance our historical understanding very far, it does have considerable applied consequences for bible translators in conditions of a dialect chain.

Blood's reconstruction is significant for the important methodological advances it illustrates. In his conclusion, Blood emphasizes the importance of using sufficient data which is available in a reliable phonemic script. Also, he stresses that:

There is a need for distributional statements, particularly of vowels in relation to finals, and of permitted consonant clusters (including complex unit phonemes) in the syllable initial position. It is quite important to know what is permitted and what is not permitted in a given environment in a given language. It is also important on occasion to know whether a given sequence

is common or rare. (pp. 107–8)

With this statement Blood is clearly laying out methodological imperatives that had to be incorporated into comparative work if it was to progress beyond the backward state it enjoyed in Mon-Khmer linguistics at the time. The importance of reliable phonemic transcription, analysis of word structure, and the compilation of comprehensive data is now clearly spelled out, and demonstrated to some extent. However, Blood's reconstruction cannot be considered to be a demonstration of the comparative method — rather it is an explanation of various *preliminary steps* that must precede comparative work with Mon-Khmer (and other) languages.

Blood assembles 428 Proto-Mnong forms, and in addition there are a number of Proto-South-Bahnaric reconstructed forms suggested. Typically the reconstructed forms coincide with their reflexes in the Preh and Bunor dialects (East Mnong), which appear to be the most conservative. The inventory of Proto-Mnong phonemes is reconstructed as follows:

*p	*t	*c	*k	*ʔ
*b	*d	*ɟ	*g	
*ʔb	*ʔd	*ʔɟ		
*m	*n	*ɲ	*ŋ	
*w	*l, *r	*j, *s	*h	
*i	*u	*iː	*uː	*iə *uə
	*ə	*eː	*əː	*oː
*a	*ɔ	*ɛː	*aː	*ɔː

The above inventory is very typical of South Bahnaric languages. Blood explains that:

The reconstruction of the consonants and consonant clusters was quite straightforward, usually being simply a matter of selecting the commonest form as the proto-form. (p. 11)

The proto-vowels are not reconstructed on the basis of a system of regular correspondences; rather Blood specifies that they are determined on the basis of "a general examination of the data" (p. 12). The proto-vowel in each case is reconstructed as the vowel which occurs in the majority of forms in a given correspondence, and vowels that occur in a minority of cases are eliminated from the proto-system. Interestingly Blood explicitly rules out the reconstruction of *ɨ on this basis (p. 12), but in the actual list of reconstructions gives some proto-forms with *ɨ, although these have parallels in Radê (a Chamic language). The above procedures do not constitute the comparative method as it is commonly

understood — rather they constitute a simple averaging exercise.

Blood does make an important contribution to the field of Bahnaric historical phonology — here is the beginning in print of a new tradition, born out of thorough training in field methods and phonemics that Blood and his fellow scholars had received. It marks a break with the practice of many traditionally trained historical linguists, especially Indo-Europeanists, who are often not accustomed to collecting lexical data directly from informants, and often work with forms taken directly as they appear in written records, which characterized previous comparative work in this field.

4. Thomas & Smith (1967): Proto-Jeh-Halang

Jeh and Halang are two very closely related North Bahnaric languages, lying adjacent to each other along an area which includes the confluence of the borders of Vietnam, Cambodia, and the Lao PDR. The geographical closeness is important, as Thomas & Smith, like Blood, emphasize the idea that reconstruction should concentrate on closely related languages:

...the study is significant in that it demonstrates the possibility of reconstruction of vowels in Mon-Khmer language comparisons, provided one starts with languages closely enough related and has enough data on dialect trends; wider comparisons have almost invariably had trouble establishing anything regular on vowel comparisons. (p. 175)

The above sentiment also echoes Blood's call for methodological rigor, particularly in relation to the investigation of vowels. The paper also has other features in common with Blood's — the work is collaborative, the data collected, prepared, and analyzed by the authors. And the method of reconstruction is conducted similarly, although in this case only two languages are treated.

The paper begins with a brief description of Jeh and Halang. The general word structure is stated ([Cə]CVC), the phonemes are listed, and some data is given on their frequency and distributional characteristics. As with the Mnong languages, the consonant systems do not differ, so the same is proposed for the proto-system. However the vowel systems do contrast, and this is where the methodology is particularly important. Thomas & Smith's first assumption is that the proto-system will be similar to that of the witness languages, so they posit as few changes as can possibly account for the development of the systems from the common source. The vowel correspondences are tabled and proto-forms are suggested on the following basis:

- where both languages show the same vowel, the same vowel is reconstructed
 - where vowels differ, the simplest development is chosen
 - where either choice is reasonable, long vowels are favored over short vowels, and breathy register vowels are favored over tense register vowels.
- According to these principles proto-phoneme assignments were made and approximately 500 Proto-Jeh-Halang etyma reconstructed.

The Proto-Jeh-Halang vowel system is as follows:

Tense Short		Tense Long		Lax Short		Lax Long	
		ia	ua			ia	ua
i	u	i:	u:	i	u	i:	u:
e	o	e:	o:	ə		e:	o:
a		a:				[a:]	

This system is virtually the same as the modern languages, except that lax short /e, a, o/ and lax long /a:/ do not occur in native words, as they are typically found in loans. Again we suggest that this is effectively just an averaging exercise, and not necessarily a reflection of historical development. The main problem with the approach of Thomas & Smith is that it essentially repeats the method Blood used for his Proto-Mnong — common phonemes are identified, and correspondences which involve different sounds are dealt with by positing one of the witness phonemes as the proto-form. The approach does not offer an historical explanation of the origins of the systems, and cannot be relied upon to determine the common ancestor to the witness languages.

5. Smith (1972): A phonological reconstruction of Proto-North-Bahnaric

Published as a short monograph, Smith (1972) has come to be regarded as the standard reference in Bahnaric historical phonology, and continues to be cited for its reconstructions.

The project is ambitious, as it aims at a reconstruction of Proto-North-Bahnaric, the oldest and most heterogeneous sub-group of the Bahnaric family. As such it should be rather close to a reconstruction of Proto-Bahnaric. Smith himself anticipates this with the statement that “the PNB reconstructed in this paper may in actuality be something like Proto-Eastern-Bahnaric or Proto-Bahnaric” (p. 11).

Five languages are used for the lexical and phonological correspondences: Bahnar, Proto-Jeh-Halang (Thomas & Smith 1967), Early Sedang (essentially the Sedang Central Dialect), Hre, and Proto-Hre-Sedang (Smith’s own reconstruc-

tion, not separately published⁴). The text begins with a crucial methodological statement:

It has been found both here and in other studies of the Mon-Khmer languages of South-Vietnam that the vowels are difficult to determine apart from the final consonant. For this reason many analysts construct "rhyming dictionaries" during their initial phonological study. Some final consonants occur with only a limited number of vowels; few, if any, occur with all vowels (Thomas 1966). In this paper all reconstructed vowels and final consonants are brought together so that the overall effect on the vowels by a given final consonant can be clearly seen. (p. 4)

This statement reflects a crucial step forward in thinking. It recognizes that the rime (root vowel plus terminal consonant) in these languages behaves as a phonological unit, and must be treated as such for historical phonological study.

Unfortunately, Smith does not take full advantage of this principle in his study. After making the above statement he proceeds to review the phonological systems of the languages in question in terms of inventories of phonemes, in the same way as his colleagues do. The inventories are surveyed, and Smith proposes a phoneme inventory from which the daughter inventories can be derived by means of the fewest changes. The problem, as we see it, is that the approach does not really treat the languages as systems, that is in terms of how the components affect each other. An understanding of this principle is hinted at in the brief mention of the importance of considering vowels in terms of the rimes in which they occur, but the follow-through is not seen. We contend that Smith should have considered the distribution of rimes within the daughter languages first before proceeding to the comparative part of the work.

The need to give the vowels special attention is particularly great in the present case because of the differences in the daughter systems — Bahnar has a 3 × 3 vowel system typical of South Bahnaric languages (with contrastive length and no vowel registers), whereas the other languages in the study have vowel registers, numerous diphthongs, and significant discrepancies between the inventories of short and long vowels. One could simply devise a proto-system which has sufficient phonemic oppositions that any of the daughters can be derived from it, but the question of the character of the sound system remains — in this case particularly whether vowel registers should be reconstructed. In common with his colleagues in the SIL tradition, Smith does not reconstruct the proto-system based upon the system of correspondences among the daughter reflexes; instead the system is decided in advance of this. The problem is instead formulated in terms of which of the daughter systems is more archaic, and this is tested by comparison with systems of other Mon-Khmer languages.

Appendix 1 of Smith compares four non-Bahnaric Mon-Khmer languages (Kuy, Bru, Khmer, and Mon) with his Proto-North-Bahnaric. Smith finds that registers correspond with North-Bahnaric in 105 lexical comparisons out of 157 sets, and considers that this correlation is sufficient to suggest that register should be reconstructed for Proto-Mon-Khmer and therefore Proto-North-Bahnaric. This leads Smith to reconstruct Proto-North-Bahnaric vocalism which is similar to Thomas & Smith's (1967) Proto-Jeh-Halang. Smith's Proto-North-Bahnaric vowels are as follows:

Tense Short		Tense Long		Lax Short		Lax Long	
		i:	u:	i	u	i:	u:
e	o	e:	o:	â		e:	o:
a		a:				a:	

The proposed Proto-North-Bahnaric consonant inventory is similar to modern Bahnar, and this is not surprising. Bahnar shows the greatest number of consonantal oppositions of any of the languages, so one need go no further for a system from which all others can be derived without positing a single innovation (only losses).

Without examining a single North-Bahnaric correspondence set, we now have inventories of proto-consonants and proto-vowels. Given some knowledge of the tendencies by which some vowels will occur with some terminals and not with others, one can now lay out the set of proto-rimes that would have been possible in the proto-language. This is now a list of predictions which can be tested: sets of corresponding daughter rimes are assembled against the proto-rimes predicted by the model. This is precisely what Smith has done, and we have the curious situation that while Smith has successfully assembled a substantial body of cognates, and discovered the regular correspondences among them, he has not used their evidence to reveal the proto-system. Rather he has simply assigned each correspondence a value according to the system he has already predicted.

Thus the major contribution that Smith makes is the presentation of 571 sets of North Bahnaric lexical comparisons, a tremendous source of data for subsequent research. However, it suffers from similar problems to other work of the same period discussed above — the proto-system is prefigured without reference to the system of correspondences among daughter forms.

6. Conclusion

We hope to have shown that the three works reviewed above do not constitute historical reconstructions, and thus must be used with extreme caution. The authors at all times acted in good faith and in perfect agreement with their training, but what they understand by comparative linguistics is something utterly alien to what we (traditionally trained historical linguists) know as the comparative method. However, any comparativist working in Southeast Asia is to some extent dependent upon such sources, and is unquestionably indebted to the SIL enthusiasts for the many important advances that have made comparative linguistics in this area not only possible, but successful. The crucial role of a fearless and comprehensive historiography is thus confirmed.

Author's address

Paul Sidwell
Department of Linguistics
Australian National University
Canberra 0200, AUSTRALIA
paul.sidwell@anu.edu.au

Notes

1. Austroasiatic is the principal substrate language stock of Southeast Asia, and is a large family comparable in breadth and depth to Indo-European, so its study has tremendous importance for comparative linguistics.
2. The present writer recently discovered why he had not received a particular journal for 2 years, and why a paper approved for publication had not in fact been published — namely that he had committed the deadly and unprofessional sin of airing various views on the historical intellectual context of certain papers/monographs. Such are the slings and arrows of intellectual fortune. The details will be communicated privately upon request.
3. The phrase “the vowel problem” appears to have been coined by Thomas — it is literally the last word in the first issue of *Mon-Khmer Studies* (Thomas 1964: 162).
4. Proto-Hre-Sedang is effectively a version of Hre, and the present author's investigations lead him to believe that Sedang is historically a phonologically innovative dialect of Hre.

References

- Blood, Henry F. 1966. *A Reconstruction of Proto-Mnong*. M. A. thesis, Department of Linguistics, Indiana University.
- Grimes, Joseph E. 1995. *Language Survey Reference Guide*. Dallas: Summer Institute of Linguistics.
- Kirk, Neile & Paul Sidwell. 1998. "Editors' Preface". *Professing Koernerian Linguistics* ed. by Neile Kirk & Paul Sidwell, v–viii. Melbourne: Association for the History of Language.
- Pike, Kenneth L. 1947. *Phonemics: A technique for reducing languages to writing*. Ann Arbor: University of Michigan Press.
- Smith, Kenneth D. 1967. "A Phonological Reconstruction of Proto Central North Bahnaric". *UNDSIL 1967 Workpapers* (manuscript sighted at SIL Bangkok).
- . 1972. *A Phonological Reconstruction of Proto-North-Bahnaric*. Language Data Series. Dallas: Summer Institute of Linguistics.
- Thomas, David D. 1964. "A Survey of Austroasiatic and Mon-Khmer Comparative Studies". *Mon-Khmer-Studies* 1.149–62.

Koerner Tabula Gratulatoria

Francisco R. Adrados, Madrid, Spain
Anders Ahlqvist, Galway, Ireland
Tsutomu Akamatsu, Leeds, United Kingdom

Cynthia Allen, Canberra, Australia
Cristina Altman, São Paulo, Brazil
René Amacker, Geneva, Switzerland
Ulrich Ammon, Duisburg, Germany
Henning Andersen, Los Angeles, California, USA

Raimo Anttila, Los Angeles, California, USA

R.E. Asher, Edinburgh, United Kingdom
Sylvain Auroux, Fontenay, France

Françoise Bader, Paris, France

Philip Baldi, University Park, Pennsylvania, USA

Allan R. Bomhard, Charleston, South Carolina, USA

Jivco Boyadjiev, Sofia, Bulgaria

Bela Brogyanyi, Freiburg, Germany

Vit Bubenik, St. John's, Newfoundland, Canada

Ranko Bugarski, Belgrade, Yugoslavia

Theodora Bynon, Loughton, Essex, United Kingdom

James Bynon, Loughton, Essex, United Kingdom

Lyle Campbell, Christchurch, New Zealand

Garland H. Cannon, College Station, Texas, USA

Gerald L. Cohen, Rolla, Missouri, USA

John Colarusso, Hamilton, Ontario, Canada

Bernard Comrie, Leipzig, Germany

John R. Costello, New York, New York, USA

William Cowan, Smiths Falls, Ontario, Canada

Regna D. Darnell, London, Ontario, Canada

A. Morpurgo Davies, Oxford, United Kingdom

David Denison, Manchester, United Kingdom

Wolfgang U. Dressler, Vienna, Austria

Sheila Embleton, Toronto, Ontario, Canada

Rudolf Engler, Worb, Switzerland

Gregory M. Eramian, London, Ontario, Canada

Eli Fischer-Jørgensen, Virum, Denmark

Jacek Fisiak, Poznan, Poland

Lia Formigari, Rome, Italy

Michael Foster, Norwich, Vermont, USA

Thomas V. Gamkrelidze, Tbilisi, Republic of Georgia

Victor Golla, Arcata, California, USA

W. Terrence Gordon, Halifax, Nova Scotia, Canada

Claude L. Hagège, Paris, France

M.A.K. Halliday, Fairlight, New South Wales, Australia

Juhani Härmä, Helsinki, Finland

Alice C. Harris, Nashville, Tennessee, USA

Ruqaiya Hasan, Fairlight, New South Wales, Australia

Gerda Hassler, Potsdam, Germany

John Hewson, St. John's, Newfoundland, Canada

Charles F. Hockett, Ithaca, New York, USA

Henry M. Hoenigswald, Philadelphia, Pennsylvania, USA

Keelung Hong, San Francisco, California, USA

Werner Hüllen, Essen, Germany

- Christopher M. Hutton, Hong Kong,
People's Republic of China
Dell Hymes, Charlottesville, Virginia, USA
Kurt R. Jankowsky, Washington DC, USA
Mark Janse, Gent, Belgium
John T. Jensen, Ottawa, Ontario, Canada
Michael Job, Marburg, Germany
Brian D. Joseph, Columbus, Ohio, USA
Jeannette Joseph, Edinburgh, United Kingdom
John E. Joseph, Edinburgh, United Kingdom
Carol F. Justus, Austin, Texas, USA
Louis G. Kelly, Cambridge, United Kingdom
Joseph F. Kess, Victoria, British Columbia, Canada
Douglas Kibbee, Urbana, Illinois, USA
Neile Kirk, Parkville, Victoria, Australia
C.H. Kneepkens, Groningen, The Netherlands
Clemens Knobloch, Swisttal, Germany
Helena Kurzová, Prague, The Czech Republic
Merja Kytö, Uppsala, Sweden
Roger Lass, Rondebosch, South Africa
Winfred P. Lehmann, Austin, Texas, USA
Saul Levin, Binghamton, New York, USA
Michel Liddle, Ottawa, Ontario, Canada
Hans-Heinrich Lieb, Berlin, Germany
David W. Lightfoot, College Park, Maryland, USA
Margarita Llitas, Valladolid, Spain
Peter Maher, Chicago, Illinois, USA
Martin Maiden, Oxford, United Kingdom
†James D. McCawley, Chicago, Illinois, USA
P.H. Matthews, Cambridge, United Kingdom
D. Gary Miller, Gainesville, Florida, USA
Donka Minkova, Los Angeles, California, USA
Salikoko Mufwene, Chicago, Illinois, USA
Robert W. Murray, Calgary, Alberta, Canada
Stephen O. Murray, San Francisco, California, USA
Brigitte Nerlich, Nottingham, United Kingdom
Hans-Joseph Niederehe, Trier, Germany
Jan Noordegraaf, Amsterdam, The Netherlands
Bert Peeters, Hobart, Tasmania, Australia
Rebecca Posner, Oxford, United Kingdom
Gary D. Prideaux, Edmonton, Alberta, Canada
Martha Ratliff, Detroit, Michigan, USA
Emilio Ridruejo, Valladolid, Spain
E. Wyn Roberts, Burnaby, British Columbia, Canada
R.H. Robins, Caterham, Surrey, United Kingdom
Rosane Rocher, Philadelphia, Pennsylvania, USA
Laurent Sagart, Paris, France
Vivian Salmon, London, United Kingdom
Joe Salmons, Madison, Wisconsin, USA
Hans-Jürgen Sasse, Cologne, Germany
Aldo Scaglione, New York, New York, USA
Paul Sidwell, Canberra, Australia
Rajendra Singh, Pierrefonds, Québec, Canada
John Charles Smith, Oxford, United Kingdom
Maxim I. Stamenov, Sofia, Bulgaria
Karen Steffen Chung, Taipei, Taiwan, Republic of China
Danny D. Steinberg, Hanno Saitama, Japan
Joseph L. Subbiondo, Moraga, California, USA
Pierre Swiggers, Leuven, Belgium
Matsui Tajima, Fukuoka, Japan
Daniel J. Taylor, Appleton, Wisconsin, USA
Talbot J. Taylor, Williamsburg, Virginia, USA
Theo Vennemann, Ried, Germany
Kees Versteegh, Nijmegen, The Netherlands
Alexander Vovin, Honolulu, Hawaii, USA
Wolfgang Wölck, Buffalo, New York, USA
George Wolf, New Orleans, Louisiana, USA
Roger Wright, Liverpool, United Kingdom

Name Index

A

Aalto, Pentti, 102
 Aarsleff, Hans, xi, xii
 Abdullaev, Abdulkarim K., 282, 285
 Abel, Carl, xxi
 Aikhenvald, Alexandra Y., 103
 Andersen, Henning, 203
 Anderson, Nicolai, 101
 Andrejev, N. D., 103
 Andronov, M. S., 103
 Antonsen, Elmer H., 241, 242
 Aronoff, Mark, 189, 200
 Ascoli, Graziadio, 101
 Austen, Jane, 273

B

Bader, Françoise, 170
 Banniard, Michel, 182
 Beard, Robert, 199
 Benjamins, John, xvi
 Benveniste, Émile, 113
 Bergsland, Knut, 102
 Biese, Yrjö Moses, 274
 Blake, Robert, 182
 Blažek, Václav, 102, 103
 Bloch, Bernard, 46
 Blood, Henry F., 304, 305, 306, 307, 308
 Bloomfield, Leonard, xxiii, xxv, 4, 5, 10, 11, 12, 13, 14, 15, 31, 45, 54, 61, 67, 82, 240
 Boas, Franz, xiv, 7, 8, 9, 13, 14, 32

Bodener, William, 25
 Boethius, 240
 Bomhard, Allan R., 103, 104
 Bonfante, Giuliano, 147, 170
 Bonfante, Larissa, 147
 Bonnerjea, René, 102
 Bopp, Franz, xiii, xxii, 101, 217
 Bouda, Karl, 102
 Bréal, Michel, xxv, 203, 205, 212
 Brinton, Daniel, 9
 Brunner, Linus, 102
 Buccini, Anthony F., 244
 Bulatova, Rimma, 103
 Bursill-Hall, Geoffrey J., xvi
 Busse, Johannes, 211

C

Campbell, Lyle, 4, 5, 6, 8, 9, 10
 Cauchie, Antoine, 211
 Charlemagne, 185
 Chomsky, Noam, x, xxiii, 30, 31, 36, 82
 Cicero, 18
 Collinder, Björn, 101
 Comrie, Bernard, 280, 286
 Čop, Bojan, 102
 Corbin, Danielle, 200
 Coseriu, Eugenio, xi, xii
 Cristofani, Mauro, 167, 168, 169
 Culler, Jonathan, xi, xii
 Cuny, Albert, 102
 Curme, G. O., xxvi

D

- Damerow, Peter, 144
 Davis, Garry W., 244
 De Dardel, Robert, 176, 181, 182
 De Mauro, Tullio, xi, xii
 de Vries, Matthias, xxi
 Delbrück, Berthold, 124, 153ff.
 Delitzsch, Friedrich, 101
 Demosthenes, 18
 Derrida, Jacques, x
 Dickens, Charles, 273
 Dixon, Roland B., 8
 Dolgopolsky, Aaron B., 102, 103, 105
 Donatus, 185
 Doroszewski, Witold, xi, xiii
 Dorsey, J. Owen, 7
 Dryden, John, 268
 Du Ponceau, Peter Stephen, 5, 9
 Dufriche-Desgenettes, A., xxi,
 Durkheim, Émile, xi, xiii
 Dworkin, Stephen, 197
 Dybo, Vladimir, 102, 103

E

- Echenique, Elizondo, 176
 Edwards, Jonathan, 5
 Einstein, Albert, ix
 Englund, Robert K., 144
 Erasmus, 18

F

- Fähnrich, Heinz, 102
 Firth, John Rupert, xvi, xx, xxiii, 32, 45,
 54, 261
 Fortunatov, Filipp Fedorovich, xi
 Fought, John, 13
 Freud, Sigmund, xi, xii
 Friberg, Jøran, 144

G

- Gallatin, Albert, 5, 6, 7
 Gatschet, Albert, 7
 Georgov, Ivan, xxi

- Gilliéron, Jules, 31
 Gimbutas, Marija, 150
 Goddard, Pliny Earle, 9, 15
 Goldsmith, John, 87, 89
 Greenberg, Joseph, 6, 104, 210
 Grimm, Jacob, 217, 239

H

- Haas, Werner, 49, 50, 51
 Hall, Robert A., Jr., xx, 176, 180, 181
 Halle, Morris, 82
 Hamp, Eric P., 222
 Hanssen, Friedrich, 192
 Harrington, John Peabody, 9
 Harris, Zellig, xiv, 49
 Hashimoto, Mantarô, 289, 290
 Hegel, G. W. F., xi
 Helimskij, Eugene, 103
 Henry, Victor, xi
 Henshaw, Henry, 7
 Herman, József, 182
 Hill, Archibald A., 47
 Hjelmstev, Louis, 30
 Hockett, Charles, xiv, 12, 13, 47, 53, 54,
 60, 68
 Hodge, Carleton T., 102
 Hoenigswald, Henry, 3, 4
 Homer, 18
 Householder, Fred, 68
 Howell, Robert B., 244, 251
 Hymes, Dell, xiv, 13

I

- Illič-Svityč, Vladislav M., 102
 Imnajtšvili, David S., 282
 Ivanov, Vyacheslav Vsevolodovich,
 xxiv, 103, 150
 Iverson, Gregory K., 244, 252

J

- Jackson, Dan, 280
 Jakobson, Roman, xi, xii, xiv, 32, 37,
 49, 52, 54, 60, 212

Janda, Richard, 243, 244, 245, 246, 248,
251

Jefferson, Thomas, 5, 6

Jespersen, Otto, xxvi

Jones, Sir William, xiii, 302

Jonson, Ben, 271, 274

Joos, Martin, 46

K

Karlgren, Bernhard, 290

Kaufman, Terrence, 222

Kern, H., 154

Kerns, John C., 104

Key, Mary Ritchie, 26

Kim, Wancin, 289, 290, 291, 292

King, J. Ross P., 289

King, Robert D., 242

Kiparsky, Paul, 82, 85, 86, 87, 89, 249

Kisseberth, Charles W., 85

Kleinschmidt, Samuel, xxi

Klimov, G. A., 102

Koerner, E. F. K., xii, xiii, xiv, xv, xvi,
xvii, xviii, xix, xx, xxvi, 3, 17, 25, 77,
109, 131, 180, 217, 239, 301

Koppelman, D. H., 102

Kornilov, G., 103

Koskinen, Kalevi E. 102

Kroeber, Alfred, 8, 10

Krygier, Marcin, 241

L

Labov, William, 249

Lafaye, Pierre Benjamin, xxi

Lakoff, George, 87, 89

Lamb, Sydney, 82

Lehmann, Winfred P., 112

Levin, Saul, 102

Lévi-Strauss, Claude, x, xxiii

Linnaeus, Carl, 7

Lucchesi, Valerio, 210

Lühr, Rosemarie, 148

Lyly, John, 271

M

Macdonnell, A. A., 160

Macpherson, Ian R., 210

Malinowski, Bronislaw, 32

Malkiel, Yakov, xxv, 190, 191, 194,
195, 196, 197, 198, 199, 200

Manaster Ramer, Alexis, 103

Mancini, Marco, 168, 169

Marlowe, Christopher, 271

Martin, Samuel E., 289

Marx, Karl, x, xii

Mascaró, Joan, 86

Mason, Otis T., 7, 8

Mayrhofer, Manfred, 112

Meillet, Antoine, xi, 113, 170

Menéndez Pidal, Ramón, xxii, 190, 192,
193, 199

Menges, Karl, 102

Menninger, Karl, 147

Meyer-Lübke, Wilhelm, 193, 194, 200

Michaëlis de Vasconcelos, Carolina, 192

Miller, Roy Andrew, 102, 289, 290, 291,
293, 294, 298

Milroy, James, 177

Milton, John, 19

Mithun, Marianne, 6

Möller, Hermann, 102

Mossé, Fernand, 271

Mudrak, Oleg, 103

N

Nam, Kwangwu, 299

Nikolayev, Sergey L., 284, 286

Noreen, Adolf, xi

Norman, Jerry, 290

Notker der Deutsche, 240

O

Öhmann, Olof, 255, 261

Otfrid, 246, 247, 251

P

Palmaitis, Mikolas, 102

Pānini, 41, 82, 154, 156, 158, 159
 Passy, Paul, 20
 Paul, Hermann, xiii
 Pedersen, Holger, 102
 Penzl, Herbert, 241, 242
 Perel'muter, Ilja Aronovič, 119, 122
 Petrarch, 19
 Pickering, John, 5
 Pike, Kenneth L., 303
 Planck, Max, ix
 Plato, 18
 Pokorny, Julius, 302
 Polinsky, Maria, 280, 286
 Postal, Paul, 56, 57
 Powell, John Wesley, 5, 6, 7, 8, 9
 Priscian, 185

R

Rajabov, Ramazan, 280
 Ramée, Pierre de la, 19
 Ramsey, S. Robert, 289
 Ramus, Petrus (see Pierre de la Ramée)
 Raven, Frijthof A., 245, 246, 247, 248, 251
 Riffaterre, Michael, 261
 Robins, Robert Henry, xvi
 Robinson, Orrin W. (Rob), 243, 244, 245, 247, 248, 249, 250, 251
 Ruipérez, M. S., 161

S

Salmon, Paul B., xx
 Salmons, Joseph C., 244
 Sandfeld, Kristian, 225
 Sapir, Edward, xiv, xxiii, 3, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15, 32, 54, 67, 82
 Sasse, Werner, 289, 290, 294, 295, 296, 297, 298
 Saussure, Ferdinand de, ix, x, xii, xiii, xvi, xxi, xxii, xxiv, 26, 30, 35, 44, 54, 165
 Schlegel, Friedrich von, xxi, xxii, 217
 Schleicher, August, xiii, xxi, 180, 217

Schmalstieg, William, 170, 171
 Schuchardt, Hugo, 31
 Shakespeare, William, 268, 270, 271, 274
 Shevoroshkin, Vitaly V., 103, 105
 Sidney, Sir Philip, 268, 270, 271
 Sievers, Eduard, xxi
 Silva Corvalán, Carmen, 178
 Smith, Adam, xi
 Smith, Kenneth D., 307, 308, 309, 310
 Smith, Regina, 244
 Sommer, Ferdinand, 138, 139
 Sophocles, 18
 Speijer, J. S., xxiv, 153, 154, 156, 157, 158, 159, 160
 Spenser, Sir Philip, 268, 270, 271
 Starostin, Sergey (Sergej) A., 103, 284, 286
 Steinthal, Hermann, 9
 Stolz, Benjamin, 105
 Sturtevant, William, 7
 Subbiondo, Joseph L., xx
 Swadesh, Morris, 45
 Sweet, Henry, xxi, 101
 Szemerényi, Oswald, xxiv, 131, 132, 134, 136, 138, 139, 148, 149

T

Taine, Hippolyte, xi
 Tarde, Gabriel de, xi
 Terentjev, V. A., 103
 Thomas, David D., 305, 307, 308
 Thomason, Sarah G., 222
 Thomsen, Vilhelm, 101
 Thumb, A., 160
 Timberlake, Alan, 204, 208
 Toporov, Vladimir N., 103
 Trager, George L., 46
 Trask, R. Larry, 176
 Trnka, Bohumil, xxvi
 Trubetzkoy, Prince Nikolai S., 49, 212
 Trudgill, Peter, 178
 Tsyburskij, V. L., 103

Twaddell, W. Freeman, 46, 242

Tyler, Stephen A., 102

U

Uesson, Ants-Michael, 102

Uhlenbeck, C. C., 102

V

Varvaro, Alberto, 182

Vine, Brent, 168, 169

Visser, Frederikus Theodorus, 268, 273,
274

von der Gabelentz, Georg, xi, xiii

von Raumer, Rudolf, 101

W

Wackernagel, Jacob, 124

Wahlgren, Erik, 255, 261

Watson, Henry, 268

Wehlitz, Hermann, 211

Wheeler, Benjamin Ide, 9

Whitman, John B., 289

Whitney, William Dwight, xiii, xxiv, 9,
153, 154, 155, 156, 158, 159, 160

Williams, Roger, 5

Wilmanns, Wilhelm, 245

Wittgenstein, Ludwig, xxii

Y

Yi, Kimun, 290, 295

Yngve, Victor, 78

Yu, Changkyun, 289, 290, 292

Yu, Changton, 299

Subject Index

A

abduction, 177, 195
 abstractness, 56
 actional, 117
 active, 117, 119, 122, 126
 adjective, 282
 adstratum, 176, 281
 affectedness, 119
 affixation, 195, 198
 affixhood, 196
 Afrasian, 102, 103, 104, 105
 agent, 119, 122
 agreement, 207, 210, 280, 281, 286, 287
 Ainu, 104
 Akkadian, 21, 22, 23
 Aktionsart, 157, 160
 Albanian, 218, 219, 220, 221, 223, 224, 227, 228, 232
 Algonquian, 4, 5, 7, 12, 15
 alliterative verse, 22
 allomorphy, 57
 allophone, 81, 177
 Altaic, 102, 104, 289, 293, 294
 Americanist linguistics, xxiii
 Americanist, 3, 4
 Amerindian, 4, 5, 6, 14
 anagram, xxii
 analogical change, 196
 analogical levelling, 198
 analogical restructuring, 244
 analogy, 179, 193, 195, 197, 241, 242, 244, 245, 251
 analytic aspectual form, 159

Ancient Greek, 157, 226
 andaluz, 178
 Andic, 282
 Anglo-Saxon, 185
 anthropological linguistics, xiv, xxiii
 anthropology, 3, 35, 36
 anticausative, 116, 117, 123
 aorist, 155, 156, 157, 158, 160, 226
 applied linguistics, 32
 apposition, 256
 Arabic, 20, 22, 24, 26, 177, 184, 195
 Aramaic, 22
 arbitrariness, 42, 44, 50, 63
 archaeologist, 175
 archaism, 285
 areal linguistics, 9
 Armenian, 126, 134, 136, 149
 Aromanian, 218, 223, 225, 226, 228, 232
 Asax, 279, 282, 283, 287
 aspect, xxiv, 159, 160
 asterisk, 180
 Asturian, 178, 190, 192
 Athabaskan, 4, 7, 12, 14
 augment, 191
 Ausgang, 193, 194, 200
 Austroasiatic, xxvi, 301, 302, 303, 311
 autonomía, 185
 autonomy, 29, 30, 31, 32, 34, 35, 36, 37, 38
 auxiliary, 205, 207, 209, 223, 224, 232
 Avar, 279, 281, 282, 283, 285, 286, 287
 Aztec, 12

B

Babylonian, 147
 back counting, 147, 149
 Bahnar, 308, 309, 310
 Bahnaric, xxvi, 302, 303, 306, 307, 308
 Balkan *Sprachbund*, xxv, 218, 219, 223, 224, 226, 228, 230, 231, 233
 Baltic, 135
 Balto-Slavic, 134, 149
 Bartholomae's Law, 112
 Basel Epigram, 166
 basic word order, 182
 Basque, 176, 177, 218, 230
 Beothuk, 7
 Bezhta, 279, 283, 284, 285
 Biblical Hebrew, 24
 bilingualism, xxv, 176, 177, 178, 184, 233
 binariness, 63
Bindevokal, 246, 247, 251
 biology, 7
 biuniqueness, 49, 56
 blocking environment, 242, 244, 248, 251
 Bloomfieldian, xxiv, 45, 303
 Boasian anthropology, 4
 borrowing, 8, 281
 Bosnian, 231
 Botorrita, 166
 Brahmin, 20
 Bru, 310
 Bulgarian, 157, 218, 225, 232

C

calque, 221
 capital, 18, 19
 cardinality, 135
 case, 115
 Castilian, 177, 178, 190, 211, 224, 232
 Castor and Pollux dedication, 169
 Catalan, 207, 211
 Caucasian, xxvi, 279
 causative, 116, 117

Cebari, 279, 283, 284, 287
 Celtic, 123, 147, 166
 Central Algonquian, 4, 11, 12, 13
 Chadic, 26
 Chamic, 306
 change, 177
 Chicago School of Sociology, 11
 Chinese, 185, 290, 295, 298
 Chippewyan, 12
 Chomskyan revolution, ix
 Chukchi-Kamchatkan, 103, 104
 Chuvash-Turkic, 104
 CILT, xviii, xix, xxv
 class prefix, 280, 282
 class suffix, 282
 class, 282, 285, 287
 classification, 6, 7, 10
 clitic, 207, 228
 closure, 69, 70, 71, 72, 73, 77
 coefficient sonantique, 165
 cognitive linguistics, 32
 cognitive phonology, 87, 89
 cognitive psychology, 36
 cohesion, 255
 collection unit, 138, 139, 140, 145, 147, 148, 149
 Colombian Spanish, 205
 comparative linguistics, 14
 comparative method, 8, 10, 11, 12, 13, 179, 183, 217, 306
 comparative reconstruction, 11
 competence, 30
 complementary distribution, 55, 61
 compound passive gerund, 271, 272, 273
 compound past tense, 204
 compound perfect gerund, 270, 271, 272, 273
 compound perfect participle, 274
 compound tense, 223
 comprehension, 70
 constituent structure, xxiv, 67, 69, 70, 73, 77
 constraint, 84

containment, 92
 Continental Celtic, 166
 convenient fiction, 68
 convergence, 184, 218
 Cornish, 25
 Cree, 11
 Crimean Gothic, 150
 Croatian, 231
 cuneiform, 143
 Current Issues in Linguistic Theory (see
 CILT)
 cursive script, 290
 Cyrillic, 22
 Czech, 24, 117, 118, 123

D

Daco-Romanian, 218, 224, 225, 226,
 228, 231, 232, 233
 Danish, 255
 deponent, 115, 116, 123
 depth of ordering, 82
 derivational constraint, 85
Diachronica, xviii
 diachrony, xiii, xxvi, 14, 82, 109, 279,
 281, 286
 dialect, 178
 Diana, 168
 Dido, xxvi, 279
 diglossia, 184
 direct object, 121, 203
 discourse boundary, 76
 discourse unit boundary, 77
 discourse, 256
 distant (linguistic) comparison, 102
 distant (linguistic) relationship, 10, 11,
 12, 101, 105
 divergence, 184
 Dravidian, 102, 104
 drift, 177, 184, 232, 242
 dualism, 42
 Duenos, 169
 Dutch, 242

E

Early Modern English (early ModE),
 xxv, 266, 267, 271, 272, 273
 Early Romanian, 184
 Early Sedang, 308
 Eastern Mnong, 305
 Elamo-Dravidian, 104, 105
 Elsewhere Condition, 95, 249
 empiricism, 43
 empty affixes, xxv
 empty morph, 194
 enclisis, 171
 endangered language, 14
 English, 19, 21, 23, 183, 185, 249, 265
 epenthesis, 83, 90
 episode, 75
 epistemology, 34
 Eskimo, 9
 Eskimo-Aleut, 103, 104
 ethnology, 9
 Etruscan alphabet, 167
 Etruscan, 141, 147
 Eurasiatic, 104, 105
 event, 75
 evolution, 9
 Ewenki, 293
 exponential system, 136

F

faithfulness, 93, 94
 Faliscan, 166, 167
 family relationship, 5, 6
 Finnish, 204, 210
 first language acquisition, xxi
 foregrounding, 256
 fragmentation, 184, 186
 French dialects, 205
 French, 19, 21, 23, 52, 185, 193, 200,
 201, 211, 224, 225, 229, 232
 functionalist, 32
 future, 160

G

- Garigliano Bowl, 167, 169, 172
 Garigliano, 170, 171
 gender, xxvi, 279, 280, 281, 284, 285
 generative grammar, 31, 69, 243, 244
 generative phonology, xxiv, 44, 85, 242, 243
 genetic relationship, 8, 10, 12, 14
 Georgian, xxvi, 121, 279
 German, 20, 22, 192, 251
 Germanic, xxv, 22, 131, 135, 136, 138, 139, 140, 147, 148, 149, 150, 184, 185, 203, 239, 240, 241, 242, 244, 245, 256, 257, 260, 265
 gerund, xxv, 265ff.
 Gilyak, 104
 Glottalic Theory, xxiv, 111, 112, 113
 glottalized consonant, 111
 glottochronology, 184
 'God's truth', xxiv, 68, 69
 Gothic Bible, 166
 Gothic, 148, 240, 257
 grammaticalization, 121, 124
 Grassmann's Law, 112
 Greek, 17, 18, 20, 21, 22, 117, 118, 119, 123, 126, 131, 133, 134, 136, 146, 148, 149, 154, 157, 165, 191, 218, 219, 223, 224, 225, 226, 227, 228, 231, 232, 233, 257
 Greenlandic, xxi
 Grimm's Law, 112

H

- habitual, 159
 Halang, 307, 308
 harmonic phonology, 87
 Hebrew, 20, 22, 29, 82
 High German, 22
 Hindi, 20
 Hindu, 41
 Hinukh, 279, 285
 historical sociolinguistics, xxv, 177
Historiographica Linguistica, xvii

- historiography, 301, 302, 311
 Hittite, 21, 118, 123, 126, 146, 165
HL (see *Historiographica Linguistica*)
 'hocus pocus', xxiv, 68
 Hokan, 8
 Homeric Greek, 154
 Hre, 308, 311
 Humboldtian trend, xiv
 Hungarian, 257
 Hunzib, 279, 284, 285
 Hupa, 12
 Hyangka, 292, 297
 hypocoristic, 284

I

- Ibero-Romance, xxv, 176, 190, 191, 194
 ICEHL, xviii
 Icelandic, 147
 ICHL, xviii
 ICHoLS, xviii
 iconicity, 73, 122
 idealism, xxii
 identity, xxii
Iliad, 20
 Illyrian, 231
 immanence, 30
 imperfect, 155, 156, 157, 160
 impersonal passive, 117, 120
 impersonal, 123, 126
 inactive, 122, 126
 independence, 35, 37
 indirect object, 121
 Indo-European, xxiv, 4, 5, 6, 7, 9, 11, 21, 22, 102, 103, 104, 109, 115, 126, 184, 302
 inertive, 116, 117, 120, 123, 124
 inferential, 157
 inferentiality, 160
 infixation, 92
 information flow, 77
 ingenerate analysis, 250
 ingenerate umlaut, 244, 245, 249
 innovation, 177, 285, 286

integration, 35, 37, 38
 interdependence, 35
 interdialect, 178, 179
 interfix, 191, 194, 195, 196, 199
 internal reconstruction, 113
 International Conference on English
 Historical Linguistics (see ICEHL)
 International Conference on Historical
 Linguistics (see ICHL)
 International Phonetic Alphabet (see
 IPA)
 interrupting RC, 70
 intransitive, 117
 invariance, 54
 invisible hand, xi
 IPA, 20, 23, 24
 Irish, 167, 261
 isolate, 182
 Istro-Romanian, 218, 225
 Italian, 19, 117, 118, 123, 126, 183,
 205, 211, 224, 225, 232
 Italic, 166, 167, 169, 170, 171
 Italo-Romance, 192
 Item and Arrangement, 48
 Item and Process, 48

J

Japanese, 104, 291, 293
 Japanese-Korean, 104
 Japanese-Ryukyuan, 104
 Jeh, 307, 308
 juncture, 47, 48, 50, 52
Junggrammatiker (see also
 Neogrammarian), 303

K

Kartvelian, 102, 104, 105
 Kensington (Rune-)Stone, xxv, 255ff.
 Khmer, 310
 Khvarshi, 279, 282, 284
 Kidero, 279, 283
 kin terms, 284

Koguryo, 289, 290, 291, 292, 294, 295,
 298, 299
 koiné, 178, 179, 255
 Kootenay, 7
 Korean, 104, 289, 290, 292, 293, 294,
 295, 298
 Kuy, 310

L

labiovelar, 165
 language planning, 184, 185
 language split, 183
Language, 4
langue, xxii, 30, 181
 Languedocien, 205
 Lapis Satricanus, 166, 167
 Laryngeal Theory, 112, 113
 Latin American Spanish, 178
 Latin, xxv, 18, 19, 20, 21, 24, 117, 118,
 123, 131, 133, 134, 136, 146, 147,
 148, 149, 150, 166, 167, 168, 170,
 171, 177, 181, 183, 182, 184, 185,
 189, 190, 191, 193, 194, 196, 221,
 222, 257
 Latinate, 182
 Latin-Faliscan, 170
 Latinism, 257
 lexeme, 189
 lexical item, 180
 lexical phonology, 85, 87, 89, 248, 249
 lexicalization, 246
 lexicostatistics, 184
 Limousin, 205
 Linear B, 146, 165
 linearity, 56
 linearization, 51
 Linguistic Circle of Saigon, 303
 linguistic family, 6, 14
 linguistic paleontology, 105
 linguistic reconstruction, 110
 linguistic sign, 42, 44, 50
 Linguistic Society of America, 4
 linguistic stock, 7, 10

Linguistic Symposium on Romance

Languages, xxv

Lithuanian, 149

loan translation, 221

loan, 285

logic, 35

logogram, 296, 297

London School, xvi

long hundred, 138, 139, 140, 147

lowering, 83

M

Macedonian, 218, 223

Manchu, 293

Manchu-Tungus, 104

Marica, 167, 168

marked order, 73

marked structure, 74, 76, 77

marked, 53, 73, 75

markedness, xxiv, xxv, 67, 69, 73, 75,
77, 122, 203, 204, 205, 208, 209, 210,
212*masse parlante*, xxi

Mayan, 41, 147

MC (see Middle Chinese)

measure unit, 137

media tantum, 115, 116, 117, 119, 121,
123, 124

medieval Greek, 231

medieval Latin, 181

medieval speculative grammar, xvi

mediopassive, 115, 118, 126

Megleno-Romanian, 218, 223, 225, 226

memory, 70

Menomini, 82

mental processing, 69

mental representation, 69

mentalism, xxiii

MHG (see Middle High German)

Middle Chinese (MC), xxvi, 289, 290,
291, 292, 295, 296

Middle English, 265

Middle High German (MHG), 240, 241,
244, 249, 250Middle Korean (MK), 290, 291, 293,
294, 295, 296, 297, 298

Middle Mongolian, 293

Middle Turkic, 294

middle, xxiv, 115ff.

middle-passive, 119

Milanese, 209

minimal pair, 51

minuscule, 18, 19

MK (see Middle Korean)

Mnong, 304, 305, 307, 308

Modern German, 240, 243, 244

Modern Greek, 18, 226

Modern Hebrew, 82

Mokok, 279, 283, 284, 285, 287

Mon, 310

Mongolian, 104, 293

Mon-Khmer, 303, 304, 306, 307, 309,
310Morpheme Structure Condition (see
MSC)

morpheme, 189

morpholexical account, 244, 245, 247

morpholexical conditioning, 239, 241,
246, 248

morpholexicalist, 250, 252

morphologization, 248, 250

morphology, 44

morphomic, 189

morphosyntactic change, 203, 212

MSC, 84, 85

Mycenaean Greek, 146, 165

N

NAAHoLS, xviii, 3

Nagari, 20

Nahuatl, 12

Nakh-Daghestanian, 279, 280, 283, 286

nationalism, 185

Naturalness Condition (NC), 47, 55, 56,
57

- naturalness, 204
 Navajo, 12
 neo-Bloomfieldian, 45
 Neogrammarian (see also *Junggrammatiker*), xii, xiii, xxii, xxv, 13, 30, 31, 239, 240, 242, 250
 Neogrammarian revolution, xiv
 neologism, 176, 284
 Netherlandic, 251
 Netherlandic dialects, 244
 neurolinguistics, 32
 neutralization, 81
 new structuralists, ix
 nominal augment, 191
 non-interrupting RC, 70
 Norse, 261
 North American Association for the History of the Language Sciences (see NAAHoLS)
 North Bahnaric, 307, 308
 Northeast Caucasian, 279
 northern Italian dialects, 224
 Norwegian, 255
 Nostratic, xxiv, 102, 103, 104, 105, 106
 noun class, 279, 286
 number system, 41
 numerals, xxiv
- O**
 object pronoun, 228
 OC (see Old Chinese)
 Occitan, 205
 Ogham Irish, 167
 OHG (see Old High German)
 OJ (see Old Japanese)
 Old Bohuslänsk, 255
 Old Chinese (OC), xxvi, 289, 290, 291, 292, 295, 296, 298
 Old English, 22, 148, 240, 241
 Old French, 194, 211
 Old High German (OHG), xxv, 150, 239, 240, 241, 243, 245, 246, 247, 248, 249, 250, 252
 Old Icelandic, 134, 147, 150
 Old Irish, 133, 134, 136, 149
 Old Italian, 210
 Old Japanese (OJ), 291, 295
 Old Korean phonogram, 289
 Old Korean, xxvi, 289, 290, 292, 293, 294, 295, 298
 Old Norse, 257
 Old Paekche, 294
 Old Prussian, 166
 Old Saxon, 22, 148
 Old Spanish, 192, 196, 197, 210
 Old Swedish, 257
 opaque, 84
 opposition, 49
 Optimal Domains Theory (ODT), 93
 optimality theory, xxiv, 89, 95
 optimality, 208
 optionality, 59
 ordered rules, xxiv
 orthography, xxiv, 180, 181
 Oscan, 166, 168, 176
 Oscan-Umbrian, 168, 170, 171
 Osco-Umbrian, 126
 over counting, 147, 149
- P**
 Paekche, 294, 295, 297
 paradigmatic, 50
 parole, 181
 participial periphrastic formation, 159
 participle, 205, 268, 273
 particle, 155
 passive gerund, 266, 270, 271, 272
 passive, 116, 123, 124, 126
 past participle, xxv, 203, 204, 205
 past perfect, 223, 226
 past, 160
 patient, 122
 pause, 47
 pedagogical grammar, 182
 Penutian, 8
 perfect gerund, 266, 267, 270, 273

- perfect participle, 268
 - perfect passive gerund, 272, 273
 - perfect passive participle, 273
 - perfect, 115, 121, 155, 156, 157, 160, 224
 - periphrastic future, 159, 160
 - personal passive, 117
 - philosophy, 36
 - Phoenician, 22, 23, 184
 - phoneme, xxi
 - phonemic analysis, 303
 - phonemic theory, 242
 - phonemics, 48, 51, 52
 - phonetic change, xxii
 - phonetic law, 11
 - phonetic transcription, 181
 - phonetics, xxi, 53, 180
 - phonogram, xxvi, 289, 296
 - phonological conditioning, 244, 245, 246
 - phonological exception, 242
 - phonological opacity, 81
 - phonological regularity, 250
 - phonology, xxi, 45, 46, 53, 81
 - phrase structure grammar (PSG), 58
 - Phrygian, 126
 - PIE root, 111
 - PIE stop, 110
 - PIE, 109, 110, 111
 - Piedmontese Italian, 233
 - pluperfect, 154, 226
 - Polish, 204
 - Portuguese, 192, 195, 208, 225, 232
 - positivism, 31
 - post-Bloomfieldian, 47, 54
 - post-Neogrammarian, xxii
 - post-structuralism, x, 33
 - pragmatics, 32, 181
 - Prague School, xxiv, 32
 - preliterate language, 175
 - preposition, 220
 - prepositional phrase, 223
 - pre-structuralist, xxvi, 32
 - preterite, 155
 - primary umlaut, 240, 242, 244, 247, 248, 249, 250
 - privative, 95
 - proclitic, 228
 - production, 70, 77
 - progressive, 159
 - proleptic apposition, 256
 - pronunciation, 182, 183
 - prosodic morphology, 92
 - Proto-Altaic, 291
 - proto-auxiliary, 159
 - Proto-Central Algonquian, 11
 - Proto-Hre-Sedang, 308, 311
 - Proto-Indo-European (see PIE)
 - Proto-Italic, xxv, 170, 171
 - Proto-Jeh-Halang, 308, 310
 - Proto-Mnong, 305, 306, 308
 - Proto-Nakh-Daghestanian, 283, 286, 287
 - Proto-North-Bahnaric, 308, 310
 - Proto-Romance, 179, 180, 181, 182, 183
 - Proto-Sardinian, 179
 - Proto-South-Bahnaric, 306
 - Proto-Tsezic, 284, 285, 287
 - pseudoreflexive, 117
 - psycholinguistics, xxiii, 32, 34, 36, 38
 - psychological reality, 68, 69
 - psychology, 35, 36
- Q**
- quasi-universal, 184
- R**
- Radê, 306
 - ranked constraints, xxiv
 - real world, 176
 - reanalysis, 177
 - reciprocal, 116
 - reconstruction, xxv, 110, 113, 175, 302, 303, 304
 - recoverability, 206
 - recursion, 58, 60, 62, 63
 - redundancy, 257
 - reduplication, 92, 93

- reflexiva tantum*, 121
 reflexive, 116, 118, 120, 121, 122, 123, 126
 regular sound change, 239, 241, 242, 245
 relative clause (RC), 70, 71
 residue, 240, 250
 resultative, 155
 retrospective, 159
 reverse, 117
 revolution (see also Chomskyan revolution, Neogrammarian revolution), xiii, xxiii
 Rhaeto-Romance, 207
 Ritwan, 4, 12, 14, 15
 Roman numeral, 141, 142, 147, 148, 149, 150
 Romance, xxv, 175ff., 189, 190, 191, 203, 207, 208, 210, 212, 221, 224, 225, 226, 228, 230, 231, 257
 Romanian (see also Rumanian), xxv, 217 ff.
 Romanian-Albanian, 222
 Romans, 146, 147
 Romany, 218
Rückumlaut, 244, 245, 246, 248, 249
 rule ordering, 82
 Rumanian (see also Romanian), 208
 Russian, 22, 134, 204, 284

S
 Sabellian, 169
 Sagada, 279, 285
 Salamancan, 190
 Sanskrit, xxiv, 20, 21, 24, 147, 153ff., 167
 Sardinian, 207
 script, 290
scripta, 185
 seal script, 290
 secondary umlaut, 240, 241, 242, 247
 Sedang, 308, 311
 semantic association, 280
 semantic change, 181
 semantics, 181
 semiolinguistics, 36
 semiology, 35
 semiotics, 36
 Semitic, 22, 23, 103
 separationist model, 192, 198
 sequential ordering, 82
 Serbian, 231
 Serbo-Croatian, 218
 sexagesimal system, 138, 140, 141
 shortening, 83
 SIL (see Summer Institute of Linguistics)
 Silla, 289, 290, 292
 simplicity, 84
 simultaneously applied rules, 82
 Sino-Korean, 292, 299
 Slavic, 135, 136, 147, 225
 sociolinguistics, xxii, 32, 34, 36, 38, 186
 sociology, 35, 36
 sound change, 5, 10, 11, 240, 243
 sound correspondences, 7
 South Bahnaric, 306, 309
 Spanish, 23, 24, 177, 178, 183, 189, 190, 195, 208, 224, 232, 257
 spelling, 180
 spelling-pronunciation, 185
Sprachbund, 230, 231 (see also Balkan *Sprachbund*)
 standard French, 205
 standard German, 244
 stock, 7
 stratificationalism, 87
 Strict Cycle Condition, 95
 strict cyclicity, 86
 structural, 186
 structuralism, ix, x, xiii, xiv, xxii, xxiii, xxiv, 160, 176, 184
 structuralist, xxv, 13, 14, 30, 31, 32, 33, 34, 44, 45, 48, 67, 69, 77, 177, 182, 183, 197, 242, 244, 251
 styleshifting, 179

stylistic device, 256
 subject, 119, 122
 substratum, 176, 184, 221, 222, 223,
 231, 311
 substructure, 182
 suffix, 190, 192
 suffixation, 121, 198
 suffixoid, 191, 198, 199, 200
 Sumerian, 21, 104, 140, 141, 143, 145,
 146, 150, 218
 Sumero-Babylonian, 140
 Summer Institute of Linguistics (SIL),
 303, 304, 309, 311
 superstock, 6
 suppletion, 57, 155
 Swampy Cree, 11
 Swedish, 255, 257
 syllable, 58, 62
 sympathy constraint, 94
 sympathy theory, xxiv, 94, 95
 synchrony, xiii, xxvi, 14, 31, 82, 279,
 281, 286
 syncope, 171
 syncretization, 208
 syntagmatic, 50
 syntax, xxi, 31, 181, 183
 system, xxv, 182, 183

T

Tagalog, 92
 teleology, 199, 242
tempuslehre, 154ff.
 tense, xxiv
 test case, 183
 theopoesis, 41, 42, 43, 44, 51, 64
 Thracian, 231
 time depth, 10
 Tocharian, 126, 172
 token, 135, 137, 140, 143, 144
 tone, 298
 topicalization, 256
tough-Movement, 227, 228, 232
 transcendence, 30

transparent, 83
 tree diagram, 179
 Trisyllabic Laxing, 86
 truncation, 200
 Tsez, xxvi, 279ff.
 Tsezic, 279ff.
 Turkic, 293, 294, 299
 Turkish, 157
 Tuscan, 19, 199
 typological analysis, 184
 typological comparison, 218
 typological plausibility, 110
 typology, 110, 115, 119, 186, 226, 228

U

Ugaritic, 22, 23
 Umbrian, 169
 umlaut, xxv, 239, 240, 241, 242, 243,
 244, 245, 246, 247, 248, 249, 251,
 252
 underlying representation, 56
 uniformitarian principle, 304
 uniformity constraint, 93
 Universal Grammar, 217, 219, 226, 227,
 228, 233
 unmarked order, 73
 unmarked structure, 76
 unmarked, 73, 75
 unstressed suffix, 190, 191
 Upper German, 244
 Uralic, 102, 104
 Uralic-Yukaghir, 104
 Uto-Aztekan, 4, 12, 14

V

variability, 184, 242
 variable rule, 249
 variation, xxv, 179, 185, 283, 284
 Vedic, 157
 Venetic, 126
 verbal morphology, 181
 versatility, 208
 Vietnamese, 304

vocabulary, 181
Volscian, 166
vowel harmony, 83, 93

W

West Germanic (WGmc.), 242
Western Greek alphabet, 167
Whitman's Law, 291
Wiyot, 15
word order, 182

Y

Yale school, 14
Yawelmani, 83, 85, 87, 90, 91, 93
Yurok, 15

Z

zero, 41ff.
Zuni, 7

Contents of Volume One

PART I

Before Saussure

1. Grammar and Philosophy in the Age of Comparativism
Lia Formigari
2. In the Shadow of the Language Garden
Jan Noordegraaf
3. Lafaye's Dictionnaire des synonymes in the History of Semantics
Gerda Hassler
4. Samuel Kleinschmidt as Syntactician
D. Gary Miller
5. Dufriche-Desgenettes and the Birth of the Phoneme
John E. Joseph
6. Sound Physiology in the Making: On the role of Henry Sweet (1845–1912) and Eduard Sievers (1850–1932) in the development of linguistic science
Kurt R. Jankowsky
7. First Language Acquisition and the Ontogenetic Development of Self-Consciousness in the Work of Ivan Georgov
Maxim I. Stamenov

PART II

Saussure

8. 'The People' and their Language in 19th-Century French Linguistic Thought
Douglas A. Kibbee

9. A Glance at the History of Linguistics: Saussure and historical-comparativism
George Wolf
10. Ferdinand de Saussure: The anagrams and the Cours
W. Terrence Gordon and Henry G. Schogt
11. Identity, Similarity, and Continuity: Saussure's and Wittgenstein's search for linguistic units
Brigitte Nerlich
12. Notes on Saussure's Old Irish Copybook
Anders Ahlqvist
13. Ferdinand de Saussure en Bulgarie
Jivco Boyadjiev

PART III

After Saussure

14. La concepción del cambio fonético de Ramón Menéndez Pidal
Emilio Ridruejo
15. "Das Glockenspiel des Ablauts": National tones in German linguistic publications between 1914 and 1945
Werner Hüllen
16. Linguistic Theory in Historical Perspective: A study of J. R. Firth's *The Tongues of Men*
Joseph L. Subbiondo
17. Between Structure and History: The search for the specificity and the originality of Brazilian linguistic production
Cristina Altman
18. How The Logical Structure of Linguistic Theory Didn't Get Published During the 1950s or 60s
Stephen O. Murray
19. How the Anti-Mentalistic Skeletons in Chomsky's Closet Make Psychological Fictions of his Grammars
Danny D. Steinberg

20. The Origins of Modern Japanese Psycholinguistics within the
Japanese Psychological Tradition

Joseph F. Kess and Tadao Miyamoto

Koerner Tabula Gratulatoria

Name Index

Subject Index